DOCUMENT RESUME

ED 149 969 SE 023 285

AUTHOR Swanson, James E.
TITLE The Effects of Adapting Elementary Science Instruction to Students' Entering Achievement Levels. Technical Report No. 417.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Aug 77
CONTRACT NE-C-00-3-0065
NOTE 192p.; Report from the Project on Conditions of School Learning and Instructional Strategies

EDRS PRICE MF-$0.83 HC-$10.03 Plus Postage.
DESCRIPTORS *Achievement; *Attitudes; Cognitive Measurement; *Educational Research; Elementary Education; *Elementary School Science; *Instruction; Knowledge Level; Science Education
IDENTIFIERS Research Reports; Science A Process Approach

ABSTRACT

This study investigates the effects of adapting elementary science instruction to students' entering achievement levels on three dependent variables: (1) performance on tests of science content knowledge; (2) mastery of lesson objectives; and (3) attitude toward science. Subjects were 235 sixth graders from four elementary schools in a small, mid-western industrial city. Two experimental and two control schools provided instruction on six lessons from the "Science: A Process Approach" (SAPA) curriculum over a ten week period. The experimental group was divided into three relatively homogenous groups based on pretest scores (high, middle or low). An attempt was made to secure mastery by as many students as possible in the experimental groups of the same science objectives. Instruction in control schools was non-differentiated and students were taught in heterogeneous classes. A 2x3 (treatment by group) factorial design was used. Two-way analyses of variance were performed on the data from a 90-item SAPA-content posttest. Students in the high and middle experimental groups attained higher scores on the posttest. Chi-square tests showed a greater number of students mastered the science objectives (80%) in the high and middle experimental groups. (Author/MR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED149969

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Richard Rossmiller
R&D Center

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SE

Technical Report No. 417

The Effects of Adapting Elementary Science Instruction to Students' Entering Achievement Levels

August 1977

Wisconsin Research and Development Center for Cognitive Learning

SE 023 285

Technical Report No. 417

THE EFFECTS OF ADAPTING ELEMENTARY SCIENCE INSTRUCTION TO STUDENTS' ENTERING ACHIEVEMENT LEVELS

by

James E. Swanson

Report from the Project on
Conditions of School Learning and Instructional Strategies

Herbert J. Klausmeier
Faculty Associate

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

August 1977

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in the University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. NE-C-00-3-0065

WISCONSIN RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING

MISSION

The mission of the Wisconsin Research and Development Center for Cognitive Learning is to help learners develop as rapidly and effectively as possible their potential as human beings and as contributing members of society. The R&D Center is striving to fulfill this goal by

- conducting research to discover more about how children learn
- developing improved instructional strategies, processes and materials for school administrators, teachers, and children, and
- offering assistance to educators and citizens which will help transfer the outcomes of research and development into practice

PROGRAM

The activities of the Wisconsin R&D Center are organized around one unifying theme, Individually Guided Education.

FUNDING

The Wisconsin R&D Center is supported with funds from the National Institute of Education; the Bureau of Education for the Handicapped, U.S. Office of Education; and the University of Wisconsin.

ACKNOWLEDGMENTS

I wish to thank Professor Herbert J. Klausmeier, my major professor, for his support throughout my doctoral studies and in the writing of this dissertation. I am also grateful to members of my examining committee, professors Gary Davis, Penelope Peterson, and Frederick Danner; and to my minor adviser, Professor Herbert Kliebard.

To friends and co-workers I am grateful for the personal and professional relationships that have made my work meaningful and my responsibilities pleasurable. Special thanks are given to Tom Sipple, who has been a constant and dependable source of intellectual stimulation and personal concern.

Further thanks are due to Suzy Gregg for her exceptional care in typing the manuscript, to Larry Winkler for his aid in the analysis of data, and to Harold Haag for his support and help in the conduct of the study.

Finally, I lovingly acknowledge those on whom I have depended the most: Karen, Daniel, Betty, Vern and Elmer.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.	iv
LIST OF TABLES	vii
LIST OF FIGURES.	ix
ABSTRACT	x
CHAPTER	
I. INTRODUCTION	1
Background of the Problem to be Investigated . .	2
Purpose and Hypotheses.	6
Method	8
Significance of the Study.	11
II. REVIEW OF LITERATURE	13
Mastery Learning Formulations and Research . . .	13
Instructional Programming for the Individual Student.	34
Components of Instructional Programming in the Present Study	46
III. METHOD	57
Description of Sample	57
Instructional Materials	61
Instructional Procedures	67
Data Collection	77
Purpose and Hypotheses.	89
Experimental Design and Data Analyses.	92
IV. RESULTS.	95
Pretest Achievement Data	95
Posttest Achievement Data.	97
Posttest Mastery Data	111
Posttest Affective Data	114

TABLE OF CONTENTS (cont.)

		Page
V.	DISCUSSION AND CONCLUSIONS	123
	Experimental Effects: Science Achievement	123
	Experimental Effects: Mastery of Objectives	130
	Experimental Effects: Attitudes Toward Science.	131
	Limitations of the Study	132
	Conclusions and Implications.	133
REFERENCES		137
APPENDICES		
A.	SAPA-Content Tests	151
B.	Notes to Teachers and Summary of Instructional Procedures to be Carried Out in Experimental and Control Schools.	153
C.	Science Attitude Inventory	157
D.	Pretest Numbers, Means and Standard Deviations for Each SAPA-Content Test	161
E.	Posttest Numbers, Means and Standard Deviations for Each SAPA-Content Test	169
F.	Numbers of Students Attaining Mastery for Each SAPA-Content Lesson.	177

LIST OF TABLES

Table		Page
1	Number of Subjects in Each Group at Each of the Four Participating Schools.	60
2	SAPA Science Lessons Taught at Level Six During Second Semester of Instruction.	63
3	Psychometric Characteristics of the Six SAPA-Content Tests Administered to Experimental and Control Groups (Posttest)	68
4	Summary of Instructional Conditions for Experimental and Control Treatments	70
5	Psychometric Characteristics of the Science Attitude Inventory (Posttest)	81
6	Summary of Time Spent on Instruction in Each Experimental and Control Classroom	83
7	Design of the Study	93
8	Regression Slopes by Group Within Each Experimental and Control School	93
9	Numbers, Unadjusted Means, and Standard Deviations for Experimental and Control Groups on Combined SAPA-Content Tests L-Q: 60-Item Pretest .	96
10	Numbers, Unadjusted Means, and Standard Deviations for Experimental and Control Groups on Combined SAPA-Content Tests L-Q: 90-Item Posttest .	99
11	Results of Two-Way Analysis of Variance on 90-Item SAPA-Content Posttest for School E_1 Paired with School C_1.	100

LIST OF TABLES (cont.)

Table		Page
12	Results of Two-Way Analysis of Variance on 90-Item SAPA-Content Posttest for School E_2 Paired with School C_2	103
13	Results of Two-Way Analysis of Variance on 90-Item SAPA-Content Posttest for Combined Experimental Schools E_1E_2 With Combined Control Schools C_1C_2	106
14	Numbers of Students at Mastery and Non-Mastery and Results of Chi-Square Tests by Group: SAPA-Content Posttests L-Q Combined	112
15	Number and Percentage of Students in Each School Attaining Mastery (80% Level) on Individual SAPA Units of Instruction	115
16	Means and Standard Deviations for Experimental and Control Schools on Science Attitude Inventory Posttest	117
17	Results of One-Way Analysis of Variance on Science Attitude Inventory Posttest for School E_1 Paired With School C_1	119
18	Results of One-Way Analysis of Variance on Science Attitude Inventory Posttest for School E_2 Paired With School C_2	120
19	Results of One-Way Analysis of Variance on Science Attitude Inventory Posttest for Combined Experimental Schools E_1E_2 Paired with Combined Control Schools C_1C_2	121

LIST OF FIGURES

Figure		Page
1	Sequential Steps of Bloom's Strategy.	22
2	Instructional Programming Model for the Individual Student in IGE.	36
3	Patterns of Objectives, Criteria of Attainment, and Sequencing	39
4	Sequence of SAPA Lesson Presentation for High, Middle, and Low Groups in Experimental Schools . .	73
5	Mean Scores on 90-Item SAPA-Content Posttest by Group for School E_1 and School C_1.	101
6	Mean Scores on 90-Item SAPA-Content Posttest by Group for School E_2 and School C_2.	104
7	Mean Scores on 90-Item SAPA-Content Posttest by Group for Combined Experimental Schools E_1E_2 and Combined Control Schools C_1C_2	107

ABSTRACT

The present study investigated the effects of adapting elementary science instruction to students' entering achievement levels on three dependent variables: performance on tests of science content knowledge, mastery of lesson objectives, and attitude toward science. Subjects were 235 children at level six (in their sixth year of schooling) drawn from four elementary schools in a small, mid-western, industrial city.

Two experimental and two control schools provided instruction on six lessons from the Science: A Process Approach (SAPA) curriculum over a period of approximately 10 weeks. In the experimental condition, instruction was organized according to a pattern of instructional programming for the individual student; curricular content and materials were adapted to the needs of three relatively homogeneous groups, formed on the basis of pretest scores (high, middle, or low entering level of achievement). An attempt was made in the experimental schools to secure mastery of the same science objectives by as many students as possible within the three groups (high, middle, low), and time and instructional techniques were varied depending upon the needs of the students in the group. For control Ss the level of pretest score was used only for the purpose of data

analysis. Instruction in control schools was non-differentiated and students were taught in heterogeneous classes.

A 2 x 3 (treatment by group) factorial design was used: levels of the treatment factor were "experimental" and "control"; levels of the group factor were "high," "middle," and "low." Two-way analyses of variance were performed on the data from a 90-item SAPA-content posttest. In general, students belonging to the high and middle groups in the experimental condition attained significantly higher scores on the achievement posttest. There was no significant difference in achievement between low students receiving differentiated instruction in relatively homogeneous groups (experimental condition) and low students receiving non-differentiated instruction in relatively heterogeneous classrooms (control condition). The same pattern of results was obtained when chi-square tests were performed on the number of students attaining mastery of lesson objectives: a significantly greater number of Ss mastered the science objectives (80% criterion) in the high and middle experimental groups, while for the low group there was no difference in the number of Ss at mastery between experimental and control conditions. On an affective inventory of pupils' attitudes toward science no differences were found between experimental and control schools.

In addition to significant main effects for treatment and group, a significant treatment by group interaction was found; the conditions associated with the interaction were discussed.

I

INTRODUCTION

The primary job of the educational psychologist, according to Carroll (1963), is "to develop and apply knowledge concerning why pupils succeed or fail in their learning at school, and to assist in the prevention of learning difficulties" (p. 723). Presently, the challenge of the educational researcher is to develop systems of education in which individuals are provided with learning experiences appropriate to their varying needs and interests, rather than being treated as "interchangeable building blocks on an educational assembly line" (Wisconsin Research and Development Center for Cognitive Learning, 1976). Individualized schooling is becoming one of the most widely practiced methods of helping children develop as rapidly and as effectively as possible their full potential as human beings and contributing members of society. Parents, teachers, and researchers recognize that students are individuals, each with certain strengths and weaknesses. An ever increasing number of schools are now adopting techniques for matching instruction to the individual student. This matching of instruction to the individual learner is a primary goal of Individually Guided Education (IGE) and a focus of the research and development conducted at the Wisconsin R & D Center for Cognitive Learning.

BACKGROUND OF THE PROBLEM TO BE INVESTIGATED

As mentioned above, a basic unifying theme in the program of research and development carried on at the Wisconsin Research and Development Center is a concern for individual differences among children and the implication of these differences for effective instruction. Programmatic research on conceptual learning and development and instructional programming for the individual student has been conducted to clarify conditions within the learner and within the learning environment that facilitate effective instruction of the individual student. In an effort to discover variables that are relevant to classroom learning, research has been conducted in classroom settings using actual subject matter concepts (e.g., Feldman, 1972; Frayer, 1970; Klausmeier & Feldman, 1975: Klausmeier, Schilling, & Feldman, 1976; McMurray, Bernard, & Klausmeier, 1975; Schilling, 1976; Swanson, 1972; Tennyson, Wooley, & Merrill, 1972). Work on the effectiveness of certain types of printed materials in teaching concepts has been done by Bernard (1975), Katz, (1976), Schilling, Bull, Feldman, and Klausmeier (1976), and McMurray, Bernard, and Klausmeier (1975). The above studies provided important information regarding instructional variables that influence learning of subject matter concepts.

In order to learn how to use the continuously accumulating knowledge concerning conceptual learning and development as well as instructional programming for the individual student in school

settings, a three-year longitudinal intervention study was begun in 1974 in which children's learning of the process-concepts of elementary science was investigated. The purposes of the study were: (1) to determine the effects of specially prepared lessons on students' learning of basic process-concepts of science; (2) to determine the combined effects of specially prepared lessons and of instructional programming for the individual student on students' achievement in a published science curriculum, i.e., Science: A Process Approach (American Association for the Advancement of Science, 1967); and (3) to determine the effects of the combined treatments on students' attitudes toward their science instruction. The intervention consisted of two main procedures. First, printed lessons were devised to teach the key process-concepts of science included in Science: A Process Approach (SAPA). These process-concept lessons were used to supplement the regular instruction on SAPA content. Second, regular instructional procedures were modified from group instruction of heterogeneous groups to a beginning level of differentiated instruction that follows a model of instructional programming for the individual student (Klausmeier, 1976, 1977). The strategy of the longitudinal intervention study was to make minimal changes in instructional procedures during the first year, then depending upon the results obtained, to move to more fully differentiated instruction during each successive semester of the study.

Results of the first year of the intervention study were

reported by Klausmeier, DiLuzio, and Brumet (1976). Results of the second year (1975-76) of the study, reported by Klausmeier, Sipple, Swanson, and Schilling (1976), indicated significantly better performance in science for those children receiving differentiated instruction in accordance with the model of instructional programming for the individual student. The differentiated instruction afforded to children in experimental IGE schools was limited, however, by the experimental design which specified that all children in the control and experimental schools proceed through the science curriculum at the same rate. Although the results obtained during the second year of the study proved both significant and positive, the changes made in instruction were still relatively minor and the time constraint did not permit near-mastery students to move to mastery under additional instruction. Enrichment was provided to students who mastered objectives early, rather than allowing them to progress through the curriculum at a more rapid pace.

During the third year of the intervention study (1976-77) instruction was arranged to permit teachers to carry out instructional programming for the individual students more fully and without a time constraint. Three small-scale studies were conducted to determine the effectiveness of three variations of the intervention procedures: (1) presentation of a process-concept lesson alone; (2) implementation of a more clearly defined instructional program for the individual student; and (3) use of the more fully differentiated

instructional program in combination with a process-concept lesson.

The present study was conducted within a framework of the longitudinal intervention study described above. In order to more clearly delineate the external conditions that facilitate concept learning by students who differ markedly in their entering levels of achievement, this study investigated the effects of one of eight patterns (Pattern 1) of instructional programming for the individual student (Klausmeier, 1977). The eight different patterns of instructional programming and the Instructional Programming Model (IPM) (Klausmeier, 1977) are described fully in chapter two of this paper (see Figures 2 and 3).

The pattern of instructional programming involved in the present study is one in which mastery of the same objectives is to be attained by all students and in which the units of instruction (in this case, the "levels" of SAPA) are to be taken in an invariant sequence. Further, learning rate is matched to the needs of the students. Inasmuch as Pattern 1 incorporates attainment of objectives to a criterion of mastery, this study also involves the application of mastery learning principles within the total program of differentiated instruction. Such a total program as that practiced in IGE schools is broader in scope than the more limited mastery approaches espoused by Block (1974), Block and Anderson (1975), or Bloom (1968, 1976).

PURPOSE AND HYPOTHESES

The present study was intended to examine the effects of a method of instruction based on instructional programming for the individual student. In this method, both differential grouping and differential pacing of students were applied to regular SAPA science instruction in the experimental treatment. Experimental Ss were placed into one of three groups (high, middle, or low) on the basis of cognitive entry characteristics (i.e., pretest level of achievement on the six SAPA lessons to be taught) and an attempt was made to secure mastery of the same science objectives by as many students as possible within each of the three groups. Teachers proceeded through a sequence of SAPA lessons at a pace determined by the needs of the group; high, middle and low classes spent only as much time as needed by the students in the group before going ahead to the next lesson in the sequence. Specifically, the purpose of the present study was to determine the combined effects of the experimental treatment (i.e., grouping students for differentiated instruction on the basis of entry level of achievement, and varying the amount of instructional time) on three dependent variables (i.e., students' science achievement, number of students attaining mastery of lesson objectives, and attitudes of students toward science). The following research questions were posed:

1. Does instruction that is organized to promote mastery of objectives in accordance with instructional programming for

the individual student facilitate children's learning of science, as measured by (a) achievement test performance, and (b) number of students who attain mastery?

2. Does instruction that is organized in accordance with instructional programming for the individual student have a significant effect upon student attitudes toward science?

In addition to the two basic research questions, there was a third question of interest to the writer:

3. Is there a relationship between the method of instruction (experimental treatment) and the instructional group to which a student belongs?

The possibility of an aptitude by treatment interaction, as expressed in question 3, influenced the choice of experimental design (i.e., 2 x 3 factorial) and analysis (i.e., two-way analysis of variance).

Hypotheses related to each of the research questions were as follows:

1a. Experimental groups receiving differentiated instruction designed to promote mastery of objectives will perform significantly better on achievement tests of SAPA-content knowledge than control groups receiving non-differentiated, large-group instruction.

1b. A significantly greater number of children will attain mastery of the science objectives in the experimental groups, taught in accordance with instructional programming for the

individual student, than in the control groups.

2. Students in experimental schools, taught in accordance with instructional programming for the individual student, will score significantly higher than control students on an inventory of attitudes toward science, indicating a more positive affect toward science.

3. Taken as the null hypothesis, there will be no significant relationship (interaction) between the method of instruction and the instructional group to which a student belongs.

METHOD

Subjects were 235 children at level six (in their sixth year of schooling) drawn from four elementary schools in a small, mid-western industrial city. A 2 x 3 factorial design was used; the factors were Differentiation of Instruction according to Pattern 1 of instructional programming (levels = present, absent), and Cognitive Entry Level based on pretest performance (levels = high, middle, low). Six units of instruction (lessons) from the published science curriculum Science: A Process Approach (SAPA) were taught to all subjects over a period of approximately ten weeks.

The study was conducted within the framework of a longitudinal intervention study of children's learning in science. The four participating schools (two experimental and two control) were involved in the longitudinal intervention and were matched according to socio-economic factors, then randomly assigned to treatment condition; all

four schools subscribed to the principles of Individually Guided Education. In the present study the treatment condition in the experimental schools was a method of instruction based on instructional programming for the individual student. In this method the attempt was made to secure mastery of the same science objectives by as many students as possible within the time constraints of the experiment. (It was not expected that all students in the "low" groups would attain mastery of the objectives by the end of the experiment.) Experimental Ss were formed into relatively homogeneous groups, and both the amount of time for instruction and the instructional tactics and materials were varied across the homogeneous groups. In control schools, the students were taught in heterogeneous classrooms with no attempt to differentiate instruction; for control Ss the level of pretest score (high, middle, low) was used only for the purpose of data analysis and not for instruction.

Subjects in all schools received basic instruction on the six SAPA lessons in the manner customarily used by the regular teachers. In the experimental schools, teachers were randomly assigned to teach a high, middle, or low group for the duration of the study. For experimental students, instructional time was dictated by the needs of the groups: children whose entry level of achievement was high were given the opportunity to progress faster through the sequences of lessons than children who required more extensive instruction. In other words, instruction was paced to the needs of the

homogeneous groups. Following initial instruction on a SAPA lesson, experimental Ss took a diagnostic-progress test to assess level of mastery. Results were used to organize further instruction so that non-mastery students (Ss scoring below the 80% criterion) were re-taught with an attempt to focus on what was not learned. Where possible, those students who had achieved mastery were encouraged to tutor (in pairs or in small groups) other students who performed below mastery on the diagnostic-progress test.

In the control schools, students kept their regular class assignments; no grouping for instruction was done on the basis of pretest performance. Teachers spent what they judged to be adequate time for instruction on each of the SAPA lessons. Following the teaching of each lesson, a diagnostic-progress test identical to the one used in experimental schools was administered. Children then proceeded immediately to the next SAPA lesson in the sequence, without benefit of remediation or reteaching.

The dependent measures for the study were (1) 15-item multiple-choice objective tests covering the lesson content of the six SAPA lessons, and (2) an inventory of student attitudes toward science. To answer research question #1a (see page 7), a two-way analysis of variance was performed on achievement data from the SAPA posttests (administered at the conclusion of instruction on all six lessons). To answer research question #1b, Chi-square analyses were performed on the number of Ss at mastery and non-mastery. The application of

a two-way ANOVA to posttest scores allowed the E to look for a treatment x group interaction (research question #3). To answer research question #2, a one-way analysis of variance was performed on attitude inventory posttest scores.

SIGNIFICANCE OF THE STUDY

The impact of instructional programming for the individual student has been the focus of both small- and large-scale studies. Results of three years of a longitudinal intervention study have shown that the IPM can be implemented with great success in science, as in other areas such as reading and mathematics. A similar approach to tailoring instruction to student characteristics has been suggested by Carroll (1963) in a learning strategy for mastery. Although much has been written on mastery learning strategies, scarcely any research on the topic has been performed in the school setting with existing curricular programs until recently. The present study is unique in attempting to examine the relationship between students' entering characteristics and conditions of instruction which follow from instructional programming for the individual student (Pattern 1).

II

REVIEW OF LITERATURE

In the present study, experimental IGE schools organized science instruction so that as many students as possible would master the same science objectives within the time limits of the study. The two theoretical formulations that underlie the instructional conditions of this study are: (1) mastery learning as described by Carroll and others, and (2) instructional programming for the individual student. The review of literature to be presented in this chapter will be organized under separate headings, each dealing with one of the major theoretical frameworks.

MASTERY LEARNING FORMULATIONS AND RESEARCH

Background

Underlying mastery learning theory and practice is an explicit philosophy about learning and teaching. Essentially it asserts that under appropriate instructional conditions virtually all students can and will learn well most of what they are taught (Block & Anderson, 1975). Particular elements of this philosophy are not new to Western educational thinking. Bloom (1974), for example, found the idea that all can learn and learn well in the writings of early educators such as Comenius, Pestalozzi, and Herbart. Cronbach (1971)

found the same idea reflected in the writings of the English philosopher John Locke. Mueller (1976) also stressed the fact that elements of the so-called mastery learning model are not new:

> Entry behavior, individualized instruction, reinforcement, diagnosis, specification of instructional objectives, remediation, feedback, "absolute" or criterion referenced evaluation, and alternative instructional modes are concepts which have enjoyed varying degrees of popularity among instructional specialists throughout the history of formalized education. (p. 42)

As early as 1926, Henry Morrison was applying what he called the "mastery formula" in the Laboratory School at the University of Chicago. As described in Peterson's (1972) review of learning strategies for mastery, Morrison first outlined what has since come to be known under the rubric "mastery learning":

> The learning units having been identified, the next problem is the technique of pedogogical attack. Here we apply what we shall call the "mastery formula": pretest, teach, test the result, adapt procedure, teach and test again to the point of actual learning. (1926, p. 81)

About the same time, in the early 1930's, a plan of differentiated instruction was being practiced in Winnetka, Illinois by Washburne (1932). Self-pacing in academic subjects was a primary characteristic of this plan. Washburne's goal was to allow each child to master each unit of work before going on to the next, without being held back by slower children or forced forward by faster children. An early study by Thompson (1941) investigated mastery learning in seventh-grade children where an experimental group was taught mathematics under a mastery condition and a matched control group

received instruction via lecture-recitation. Thompson found that the experimental group showed greater gains on the dependent measure (Stanford Arithmetic Test) than the control group.

Although a learning strategy for mastery may be supported by the ideas of Bruner (1966), Glaser (1968), Goodlad and Anderson (1959), Skinner (1954), and Suppes (1966), it was indisputably Carroll (1963) whose work has been most influential in learning-for-mastery conceptualizations. Implicit in Carroll's model of school learning is the assumption that, given enough time, all students can conceivably attain mastery of a learning task. Rather than defining aptitude as an index of the level to which a student could learn, Carroll defined aptitude as a measure of learning rate, i.e., as a measure of the amount of time the student would require to learn under ideal conditions of instruction. A student with high aptitude for a subject would learn it quickly, while one with a low aptitude would learn it more slowly.

In its simplest form, Carroll's (1963) model of school learning proposed that if each student was allowed the time (s)he needed to learn to some criterion level, and if the necessary time were spent, then the student probably would attain the level. However, if the pupil was not allowed enough time or did not spend the required time, then the degree to which (s)he would learn could be expressed as follows:

$$\text{Degree of School Learning} = f\left(\frac{\text{Time Spent}}{\text{Time Needed}}\right)$$

This equation means the degree of school learning will depend on the time the student actually spent in learning, relative to the time (s)he needed to spend.

Carroll believed that in the school situation the time spent and the time needed were influenced not only by learner characteristics but also by the characteristics of instruction. In an expanded version of the equation the time spent, or the numerator in Carroll's fraction, is influenced by two factors: perseverance (the amount of time a student is willing to spend actively engaged in learning) and opportunity to learn (the classroom time allotted to learning). The time needed, or the denominator of the fraction, is likewise determined by several factors: a student's aptitude for the subject, the quality of instruction (including the instructional materials), and the student's ability to understand instruction. If the quality of instruction was high, then the student would readily understand it and would need little time to learn the subject beyond that required by his or her aptitude; but if the quality of instruction was low, then the student would need extra time. The complete Carroll model including determinants of time spent and time needed can be represented as follows:

$$\text{Degree of School Learning} = f\left(\frac{\text{1. Perseverance} \quad \text{2. Opportunity to learn}}{\text{3. Aptitude} \quad \text{4. Quality of instruction} \quad \text{5. Ability to understand instruction}}\right)$$

This equation means the degree of school learning of a given subject depends upon the student's perseverance or the student's opportunity to learn, relative to his or her aptitude for the subject, quality of instruction, and ability to understand this instruction.

Most, but not all, goals of the school, according to Carroll, can be expressed in the form of learning tasks or a series of such tasks. By learning task Carroll means "going from ignorance of some specified fact or concept to knowledge or understanding of it, or of proceeding from incapability of performing some specified act to capability of performing it." As clarified by Carroll, to call the unit of learning a task does not necessarily imply that the learner must be aware that he is supposed to learn, although in most cases it happens that such awarenesses on the part of the learner are desirable. Summarized in the briefest of terms, Carroll's (1963) model of school learning says that the learner will succeed in learning a given task to the extent that (s)he spends the amount of time (s)he needs to learn the task.

Bloom's (1968) Learning-For-Mastery (LFM) approach evolved from the model of school learning originally formulated by Carroll. The LFM strategy is basically a group-based, teacher-paced approach to mastery instruction wherein students learn, for the most part, cooperatively with their classmates. It is designed for use in situations where the instructional time available to students in a particular subject is relatively fixed and restricted. Using Carroll's

conceptual model, Bloom reasoned that if aptitude were predictive of the "time needed," but not of the level of achievement, it should be possible to set the expected level of achievement at a criterion of mastery. Then, by matching opportunity and quality of instruction differentially, the teacher should be able to ensure a mastery level of achievement for each student.

Extending this logic, Bloom (1968) predicted that three variables could account for most of the variation in achievement: (1) student cognitive entry behaviors (the extent to which the student has already learned the basic prerequisites to the learning to be accomplished), (2) affective entry behaviors (the extent to which the student is motivated to engage in the learning process), and (3) quality of instruction (the extent to which instruction is appropriate to the learner). Bloom believes that by using a mastery learning strategy these three variables can be altered, and as a result the variation in student achievement will greatly diminish.

Some of the basic features of Bloom's LFM strategy are summarized in the following outline (adapted from Block & Burns, in press):

1. A learner must understand the nature of the task to be learned and the procedures to be followed in learning the task.
2. The formulation of specific instructional objectives for the learning task is important in the LFM approach.
3. It is helpful to break a course or subject into smaller

units of learning and to test at the end of each unit.

4. The teacher should provide feedback on the learner's specific errors and arrange for correctives (activities for reteaching) after each test.
5. Judging from student performance on the formative tests (sometimes called diagnostic-prescriptive tests), the teacher must find ways to alter the time and type of learning activities some individuals need to learn.
6. Student effort is enhanced when small groups of two or three students meet to review their test results and to help one another overcome the difficulties identified by the formative testing.

Block (1971, 1974) and more recently Block and Anderson (1975) have refined and elaborated upon Bloom's ideas so as to make them more systematic and practical. The strategy which has resulted will be referred to as the Bloom/Block model of Learning-For-Mastery (LFM). This model specifies certain preconditions as well as operating procedures for the implementation of LFM in a school setting.

Preconditions. The teacher who wishes to use an LFM approach begins by deciding what "mastery" of the subject will be. First, the teacher specifies what all students will be expected to learn by formulating a set of instructional objectives. Second, (s)he prepares a final, or "summative," examination (Bloom, Hastings, & Madaus, 1971) over these objectives and determines to what level of

mastery all students will be expected to perform. Next, the teacher breaks an entire course into a sequence of smaller learning units, where each unit typically covers about two week's time. Simple feedback/correction procedures are also developed for each unit; sets of alternative instructional materials or "correctives" are keyed to each item on the diagnostic-progress tests. According to Block and Anderson (1975), these correctives are intended to teach the content in a way that differs from initial instruction, and can include small-group study sessions, individualized tutoring, and/or alternative learning aids. With the help of appropriate correctives, the student who encounters difficulty in learning certain concepts during initial instruction can select alternative ways of learning the unmastered material.

Operating Procedures. Having met the preconditions for LFM instruction, the teacher is ready to teach. Since students may not be accustomed to learning for mastery, Block and Anderson (1975) suggest that the teacher spend some time at the outset orienting students to the LFM strategy. Following this orientation period, the teacher generally teaches the regular unit of instruction using his or her "customary group-based instructional methods." When this initial instruction is completed, the teacher administers the unit diagnostic-progress test. Instead of moving immediately to the next unit, the teacher uses the results of this test to "certify" the progress of those students achieving mastery and encourage those who

are not progressing satisfactorily (students at less-than-mastery) to use the unit's correctives to complete their learning. Block and Anderson (1975) emphasize that these latter students are given the responsibility of completing their learning of the present unit before initial instruction for the next unit can begin. The sequential steps of the Bloom/Block LFM strategy are summarized in Figure 1. Essentially, the components of this mastery model are: (1) formal specification of a comprehensive set of cognitive objectives, (2) instruction, (3) frequent formative/diagnostic evaluation, (4) corrective or remedial instructional treatments, and (5) criterion referenced summative evaluation.

An indictment of the Bloom/Block LFM strategy has been voiced by Mueller in regard to the operating procedure in which non-mastery students must complete learning before instruction on the next unit can commence: "Under the mastery model the students who attain the prespecified instructional objectives first must, in a sense, 'wait around' for their slower learning classmates" (1976, p. 45). This is a definite handicap for the teacher who wishes to adapt both rates and methods of instruction to the characteristics of individual students. In a school using the LFM approach, even though a significant portion of a heterogeneous class may have mastered a given unit of instruction, no one can begin the next unit until the non-masters have taken the correctives and demonstrated subsequent learning of the material. On the other hand, instructional program-

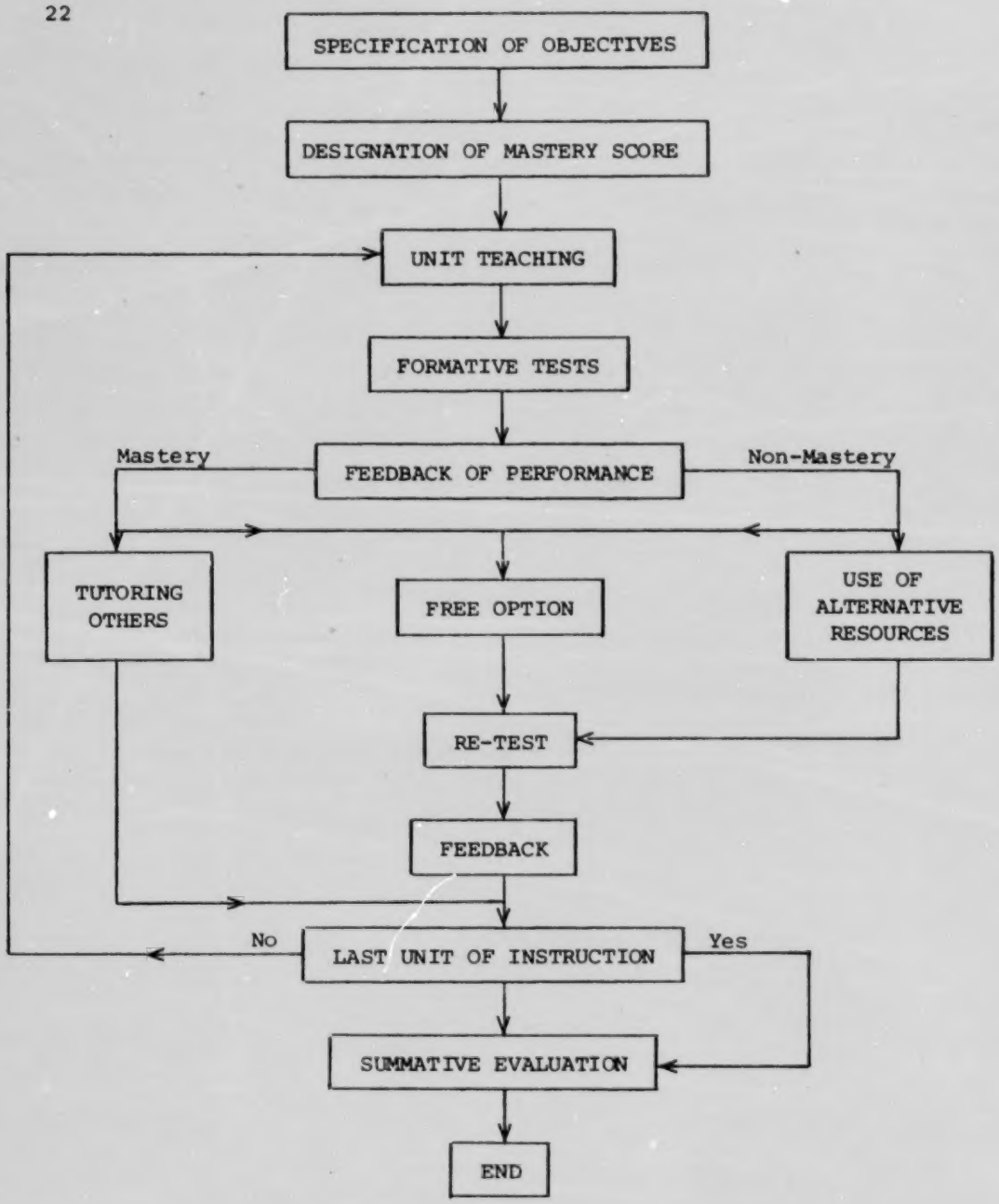


Figure 1. Sequential Steps of Bloom's Strategy (from Kersh, 1971, p. 3)

ming for the individual student (Klausmeier, 1977) permits differential pacing of students who are grouped according to entry levels of achievement. While in the Bloom/Block formalization of mastery principles there is no provision for pretesting prior to the onset of instruction, a more sophisticated application of mastery principles, such as instructional programming for the individual student, includes pretesting to identify entry levels of achievement and to pinpoint instructional needs. Another limitation of the Bloom/Block strategy is the fact that it was designed for use in age-graded, "traditional" classrooms where available instructional time is relatively fixed and inflexible. Mueller (1976) is highly critical of Bloom's contention that LFM maximizes learning for individual students:

> In traditionally organized schools, where learning is divided into grade levels and semesters, and where students are advanced together through these fixed-time instructional units, the mastery model cannot be used to maximize learning for all students. Since students learn at different rates, it is not legitimate for an instructor to expect all students to achieve the same level of performance (mastery) over the same amount of material in the same amount of time. (p. 44)

Whereas the Bloom/Block strategy is limited to age-graded schools, instructional programming for the individual student (Klausmeier, 1977) can be applied to more flexible learning situations, e.g., IGE schools.

In the present study, mastery learning principles are incorporated in a program of differentiated instruction based upon a

model of instructional programming (Klausmeier, 1976, 1977) whereby learning time is varied to match students' needs within a non-graded IGE school. Before providing a detailed description of instructional programming for the individual student and Individually Guided Education, a review of recent mastery learning research will first be presented.

Mastery-Learning Research

Many interesting claims have been made by Bloom (1976) for a mastery learning strategy. In his book, Human Characteristics and School Learning (Bloom, 1976), many empirical studies are reported to back up these claims. Peterson (1972) summarizes the major contentions of the mastery learning strategy as follows:

1. The mastery-learning strategy suggests procedures within the context of ordinary group-based instruction whereby student learning can be managed to promote fullest development.
2. With a mastery-learning strategy 75 to 90 percent of the students achieve to the same high level as the top 25 percent in a typical instructional approach.
3. All students can master what they are taught.
4. Student learning is more efficient under a mastery-learning approach.
5. Mastery learning produces greater student interest in and

attitude toward subject matter than traditional approaches. Of the studies reviewed by Peterson (1972), only one study had systematically attempted to investigate the effectiveness of different procedures. Thus, point 1 above remains largely unsubstantiated. Regarding point 2 (75 to 90 percent of the students achieve to the same high level), Peterson found only six of fourteen separate studies in which 75 to 90 percent of the students achieved the level of typical A or B students. Bloom (1976), on the other hand, writes that "in the many studies reported by Block (1971, 1974) and by Peterson (1972), there is considerable evidence that mastery learning procedures do work well in enabling about four-fifths of all students to reach a level of achievement which less than one-fifth attain under non-mastery conditions" (p. 5).

The third claim of mastery learning strategists--all students can master what they are taught--is an underlying principle of the Carroll (1963) model of school learning. This claim is based on the assumption that aptitude can be viewed as learning rate or the amount of time a pupil will need to learn a task. In theory at least, if given enough time, all students should be able to attain mastery of a given learning task. A corollary of this point is the statement that "individual differences can be eliminated." Bloom does in fact acknowledge the existence of individual differences, but he believes that they need not be reflected in school achievement. Bloom's notion that mastery learning results in vanishing individual differences

has been widely challenged (Buss & Poley, 1976). Research reviewed by Peterson (1972) suggests that individual differences cannot be eliminated.

A fourth claim associated with mastery learning is that student learning is more efficient. The validity of this claim rests upon the degree to which classroom time is used to greatest advantage. If the mastery learning strategy operated ideally, it would certainly be more efficient for the individual than a group-based approach. However, the mastery approach is not always applied as Carroll's (1963) model specifies. That is, conditions in the classroom are usually less than optimal, because students cannot always be allowed an indefinite amount of time and students do not always make the most efficient use of the time they are given.

The fifth claim of mastery learning strategists--greater student interest in and attitude toward subject matter--is based on Bloom's (1971) assumption that there is a relationship between school achievement and attitudes. The studies reviewed by Peterson (1972) suggest that a majority of students have a positive attitude toward a mastery approach to learning. It has not been empirically shown, however, that this positive attitude transfers to interest in subject matter.

A study by Shepler and Romberg (1969) abstracted in Peterson (1972) used Bloom's mastery-learning strategy to teach a unit of instruction in probability and statistics to 25 sixth-grade students

of average to above-average ability. On a 72-item test the mean pretest percentage correct was 37.9% and after instruction the posttest percentage was 92.8%. Of the 25 students, 21 (or 84%) were considered to have attained mastery (i.e., to have answered correctly 90% or more of the items), while all students answered correctly at least 80% of the items. Four weeks after the posttest, a 72-item retention test was administered to the experimental group of 25 sixth-graders (Romberg, Shepler, & King, 1970). A comparison of posttest and retention performance revealed that of the 29 students who had attained mastery initially, 17 were still above the 90% level of mastery four weeks later. Of the total experimental group, 92% of the students performed at the 80% level of mastery on the retention measure.

Also reported in Peterson (1972) was a study by King and Romberg (1970) in which a unit on mathematical proof was taught to sixth-grade students. An experimental group (mean IQ of 117.6) and a matched control group (mean IQ of 121.3) were employed; the unit of instruction was taught to the experimental group using the LFM strategy, while the control group received no instruction. On the posttest the experimental group demonstrated mastery of all unit objectives. Mean posttest score for the experimental group was significantly ($p < .0001$) greater than the mean for the control group.

Two large-scale studies by Kim (1970, 1971) were conducted in Korean schools using Bloom's mastery-learning strategy. A pilot

study was carried out in 1970 with 136 experimental subjects and 136 matched control subjects enrolled in the seventh grade. Both groups of students were taught a unit on geometric figures in eight sessions by their regular classroom teachers. The experimental students received instruction according to the mastery strategy, including formative tests, remedial "correctives," and review questions. Mastery was defined as 80% correct on the summative test. Results of the pilot study showed that 74% of the experimental group achieved mastery, while only 40% of the control group mastered the material. Kim's (1970) main study involved nine middle schools (5,800 seventh-grade students) in Korea. English and mathematics were taught for a period of eight weeks using a mastery strategy in experimental classrooms. In comparing two experimental schools with two matched control schools, it was found that 73% and 72% of the Ss in the experimental schools attained mastery in English and mathematics, respectively; 47% and 46% of the non-mastery-taught control subjects attained mastery of the same material. Kim (1970) later extended the research to include English, mathematics, physics, and biology instruction in Korean first-grade classrooms. Results (N = 25,887) indicated that 50% of the experimental Ss attained mastery in English, 48% in mathematics, 30% in physics, and 22% in biology; in non-participating control schools, 44% of the students achieved mastery in English, 26% in mathematics, 8% in physics, and 3% in biology. Yet another large-scale study by Kim (1971) applied the

mastery-learning approach to arithmetic and science instruction in 99 fifth- and sixth-grade classrooms. Non-participating, matched control schools numbered 80. Results for the fifth-graders revealed that 42% and 39% attained mastery in arithmetic and science, respectively; among the control students, 12% and 7% achieved mastery in arithmetic and science. Of the experimental sixth-graders, 45% and 46% attained mastery in arithmetic and science, respectively; of the control students, 18% and 12% attained mastery in the same subjects.

Of sixteen studies reviewed by Peterson (1972) which compared achievement of mastery-taught experimental groups with non-mastery-taught control groups, the experimental groups outperformed the control groups in fifteen studies. Only one study (Green, 1969) reported no difference between experimental and control students in achievement on the learning measure.

Block and Burns (1977) have conducted an extensive review of the extant literature on mastery learning through the summer of 1975. They utilized published research in professional books and journals, locatable unpublished or soon-to-be-published papers from professional research meetings, and unpublished doctoral dissertations. The reported studies of mastery learning were then classified according to the "research tradition" from which they derived. Block and Burns distinguish between studies belonging to the Learning-For-Mastery (LFM) tradition, and those belonging to the Personalized System of Instruction (PSI) tradition; LFM evolved from the

field of education, while PSI evolved from the field of psychology. Exemplifying the LFM tradition is Bloom (1968, 1976). The ideas of Keller (1968) led to the formalization of PSI, which is basically an individually-based, student-paced approach wherein students typically learn independently of their classmates. The PSI strategy is akin to programmed instruction and generally quite different from those mastery strategies in which students learn in various size groups, rather than independently.

In the review of mastery literature conducted by Block and Burns, four types of research are identified (according to the kinds of questions asked by each):

Type 1: Does a mastery approach to instruction really work?

Type 2: Does a mastery approach have positive affective consequences for students?

Type 3: How and why do mastery approaches to instruction work?

Type 4: How can the accumulating knowledge about mastery learning be translated into statements about the ways in which such approaches may be implemented?

Research of Types 1 and 2 are directly related to the questions addressed in the present study; and for this reason, a discussion of Type 1 and Type 2 research is provided below.

Type 1 Research. In the Block and Burns (1977) review of Type 1 studies, there are sufficient data to show that mastery-taught students exhibit greater learning, on the average, than non-

mastery-taught students. This is especially apparent in the achievement data and, to a lesser extent, in the retention data. Excluding studies of the PSI tradition, it was reported that mastery-taught groups performed significantly better than non-mastery-taught groups in 14 of the 24 studies reviewed. In another 6 studies mastery-taught groups performed better, although statistical significance was not attained. Further, Block and Burns indicate that when LFM-taught students have achieved better on summative examinations than non-mastery-taught students, they have typically exhibited 23% higher scores. Likewise, when LFM-taught students have performed better on retention measures, they have typically exhibited 28% higher scores than non-mastery-taught students. In the studies reviewed by Block and Burns, mastery-taught students have never performed poorer on the retention measures than the non-mastery-taught students.

Whereas most Type 1 mastery learning studies have utilized only a few classrooms or groups, several large-scale studies (Anderson, 1976; Jones, 1974; Kim, 1970, 1971; Kim et al., 1974; Lee et al., 1971; and Okey, 1975) have utilized many classrooms. In particular, Anderson, Jones, Kim et al., Lee et al., and Okey have compared the achievement of 12, 10, 12, 99, and 14 mastery-taught classrooms, respectively, with the achievement of 12, 10, 8, 80, and 14 non-mastery-taught classrooms. Except for the studies by Anderson and Jones, the learning data from these large-scale studies indicate that mastery approaches generally produce both significantly greater

student achievement and greater student retention across classrooms. In regard to variability in student learning, Block and Burns report that on immediate achievement tests as well as on retention tests the mastery-taught students exhibit less variability in learning than non-mastery-taught students.

Of the many LFM studies reviewed by Block and Burns, the majority were conducted at the college level; only eleven of the reported studies involved elementary (K-8) children. Okey (1975) used primary level children studying mathematics and language arts; Anderson (1976) used students in grades 1-6 studying mathematics; Kim et al. (1974) used grade two Korean children studying moral education and Korean arithmetic; Okey (1974) investigated the teaching of fractions at grades three and four; Burrows and Okey (1975) studied the teaching of geometry at grades four and five; Lee et al. (1971) examined arithmetic and science instruction at grades five and six; and Jones (1974) studied grade seven children in the area of social studies. Four studies (Anderson, 1973; Block, 1970, 1972; Fiel & Okey, 1974) were conducted at the eighth-grade level involving instruction in matrix algebra and graphs. It is important to note that in only one of the eleven studies reviewed by Block and Burns was the duration of instruction longer than four weeks. In nine of the eleven studies of elementary children, the duration of experimental treatment was only one to two weeks. Clearly, the effectiveness of mastery-learning procedures cannot be adequately assessed in such short time

intervals; there seems to be a dearth of studies at the elementary level in which mastery principles are applied for extended periods of time. In the present study, mastery procedures incorporated in Pattern 1 of instructional programming for the individual student (Klausmeier, 1977) are carried out for a period of ten weeks.

Type 2 Research. According to Block and Burns (1977), in the early 1970's some researchers, encouraged by the results of the Type 1 research, began to look beyond the effects of mastery approaches on student learning. In particular these researchers addressed the matter of affective implications. The affective consequences of mastery approaches to instruction have been studied under a wide range of conditions. The samples have been large and small, and have been drawn from primary grades through university-level classes. In addition, student "affect" has been measured with a variety of instruments. The data reviewed by Block and Burns indicate that across this range of conditions the mastery approaches have typically elicited more favorable affective responses from students than the non-mastery approaches. In particular, the mastery approaches have had a positive impact on students' interest in and attitude toward the subject matter learned, as well as on academic self-confidence, attitudes toward cooperation, attitudes toward instruction, and self-concept. Studies dealing with the affective consequences of mastery instruction have been done by Anderson (1976), Lee et al. (1971), and Poggio et al. (1975).

In summary, the Type 1 and Type 2 research provides evidence that mastery approaches to instruction do produce more desirable "products in pupil growth" (Dunkin & Biddle, 1974) than the non-mastery approaches they seek to replace. One of the distinguishing characteristics of mastery approaches to instruction has been their perceptible impact on student "products" and, especially, student achievement and retention.

INSTRUCTIONAL PROGRAMMING FOR THE INDIVIDUAL STUDENT

The Wisconsin Research and Development Center for Cognitive Learning has as its major goal the discovery, development, dissemination and implementation of concepts and principles of effective individualized schooling. Individually Guided Education (IGE) is a direct result of applying these processes to the problem of individualized schooling. IGE is a comprehensive alternative to schools organized in either age-graded, self-contained classrooms or in subject-centered departments, and it aims to provide high quality education of the total child.

> IGE aims to provide quality education that yields high student achievement, develops the abilities underlying those achievements, and contributes to healthy personality development. This aim becomes a reality in IGE schools that provide an environment in which the individual students learn at rates appropriate to each student and in a manner suitable to each student's learning style and other intellectual and personal characteristics. (Klausmeier, 1977, p. 7)

In IGE, a model of instructional programming has been conceptualized that describes how materials, activities, time, and teacher

direction are purposefully varied to ensure continuous pupil progress. The Instructional Programming Model (IPM) (Klausmeier, 1976, 1977; Klausmeier, Sorenson, & Quilling, 1971) was devised for use in IGE schools. The IPM is a design for adapting instruction to the needs of the individual students and is intended for use by IGE teachers in planning, implementing, and evaluating the instructional programs of students; the model is also intended for use by textbook writers and curriculum developers.

Instructional programming for the individual student specifically takes into account each student's beginning level of performance, rate of progress, style of learning, motivational level, and other characteristics in the context of a school's total educational program. In this regard, the "individualized instruction" afforded to children in IGE schools is more comprehensive than most nationally distributed "individualized programs." As explained by Klausmeier:

> In principle, the nationally distributed "individualized programs," which consist of identical sets of books, films, tests, and other materials to be used in the same manner by a large number of students, are in direct contradiction to the concept of instructional programming for the individual student. The many students using these materials have identical curricular programs; they do not have instructional programs specifically designed for them. (1977, p. 56)

As shown in Figure 2, there are seven major steps in the IPM. In Step 1, educational objectives are specified for the student population of a building; in Step 2, the range of objectives that

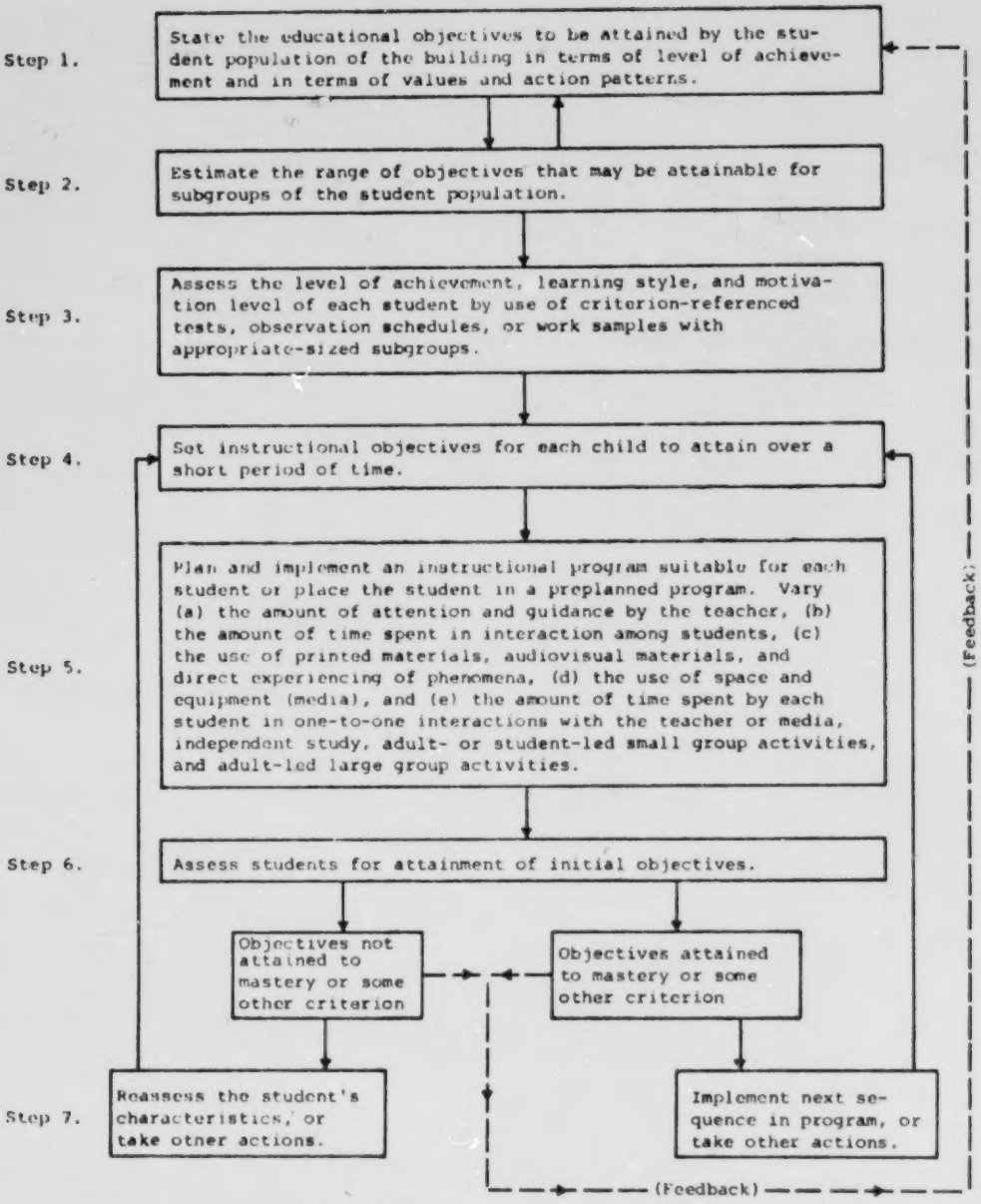

Step 1.
State the educational objectives to be attained by the student population of the building in terms of level of achievement and in terms of values and action patterns.
Step 2.
Estimate the range of objectives that may be attainable for subgroups of the student population.
Step 3.
Assess the level of achievement, learning style, and motivation level of each student by use of criterion-referenced tests, observation schedules, or work samples with appropriate-sized subgroups.
Step 4.
Set instructional objectives for each child to attain over a short period of time.
Step 5.
Plan and implement an instructional program suitable for each student or place the student in a preplanned program. Vary (a) the amount of attention and guidance by the teacher, (b) the amount of time spent in interaction among students, (c) the use of printed materials, audiovisual materials, and direct experiencing of phenomena, (d) the use of space and equipment (media), and (e) the amount of time spent by each student in one-to-one interactions with the teacher or media, independent study, adult- or student-led small group activities, and adult-led large group activities.
Step 6.
Assess students for attainment of initial objectives.
Objectives not attained to mastery or some other criterion
Objectives attained to mastery or some other criterion
Step 7.
Reassess the student's characteristics, or take other actions.
Implement next sequence in program, or take other actions.
(Feedback)
(Feedback)

Figure 2. Instructional Programming Model for the Individual Student in IGE. (Adapted from Klausmeier, Quilling, Sorenson, Way, & Glasrud, 1971, p. 19.)

may be attainable for subgroups of the student population are specified; in Step 3, an assessment (pretest) of each child's level of achievement, learning style, and motivational level is conducted; in Step 4, instructional objectives are set for each child for a specified period of time; in Step 5, an instructional program suitable for each student is planned and carried out; in Step 6, assessment (posttest) is carried out to measure attainment of initial objectives. Both Step 3 and Step 6 involve the use of criterion-referenced tests. Following Step 6, branching occurs in which further instruction is dependent on whether the specified objectives are achieved to a desired criterion set by the school. The desired criterion might be mastery by all students or achievement by each student to a level judged appropriate for the student. If the criterion is mastery and the student masters the objectives, the next sequence in the program is implemented or other actions are taken; if objectives are not mastered, the student's characteristics are reassessed or other actions taken (Klausmeier, 1976, 1977).

The steps in instructional programming outlined above are usually carried out with closely related sets of objectives during relatively short periods of time, such as three to six weeks. These short time periods of time, during which a group of objectives is related to instructional materials and activities, are considered to be units of instruction (Klausmeier, 1977). Three critical matters dealing with units of instruction are as follows: (1) whether

the objectives incorporated in the unit are to be attained by all students, (2) whether the criteria that are specified for attainment of the objectives are the same for all students, and (3) whether the units of instruction are to be taken in a fixed sequence. These three matters are handled differently both among various curricula and also within any one curricular field. Of these factors--common or variable objectives, full mastery or variable attainment of objectives, and an invariant or a variable sequence of instructional units--there are eight possible combinations. These eight patterns of instructional programming for the individual student, as shown in Figure 3, are:

1. Common objectives, full mastery, invariant sequence across units.
2. Common objectives, full mastery, variable sequence across units.
3. Common objectives, variable attainment, invariant sequence across units.
4. Common objectives, variable attainment, variable sequence across units.
5. Variable objectives, full mastery, invariant sequence across units.
6. Variable objectives, full mastery, variable sequence across units.
7. Variable objectives, variable attainment, invariant

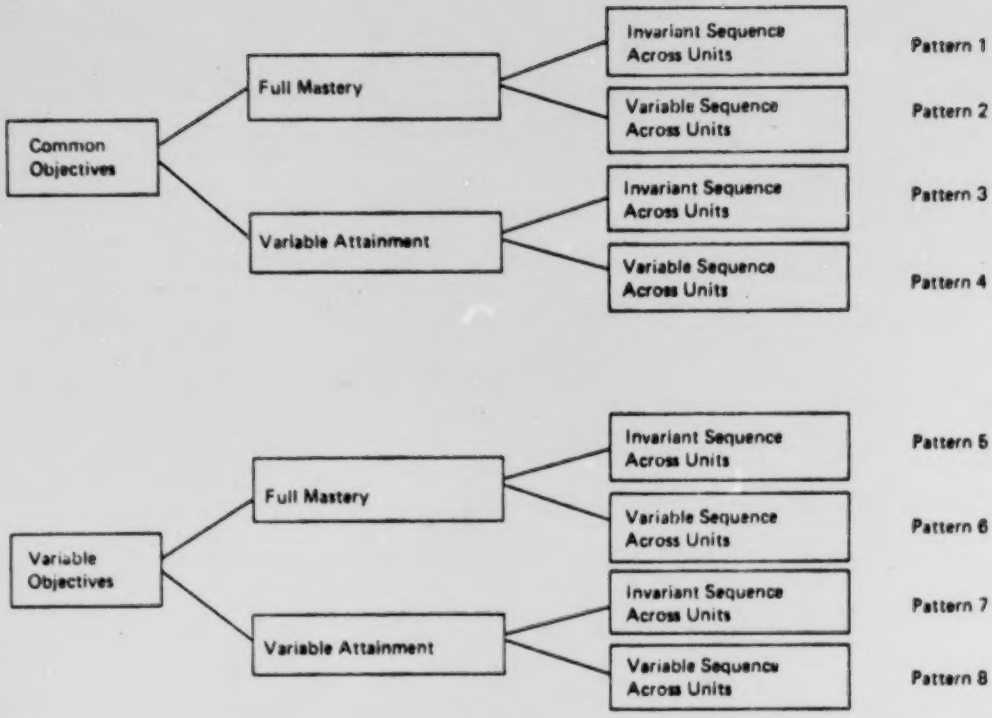


Figure 3. Patterns of objectives, criteria of attainment, and sequencing. (Taken from Klausmeier, 1977, p. 61.)

sequence across units.

8. Variable objectives, variable attainment, variable sequence across units.

Informative, detailed descriptions are available regarding instructional programming in mathematics (Romberg, 1976, 1977), in reading (Otto, 1977; Otto & Chester, 1976; Venezky & Pittelman, 1977), in social studies (Bechtol & Conte, 1976), and in science (Haney & Sorenson, 1977). Based on Klausmeier's (1977) descriptions of instructional programming in Individually Guided Elementary Education, examples of curricular programs that typify each of the eight patterns of objectives, criteria of attainment, and sequencing of units of instruction are offered below.

Pattern 1: Common objectives, full mastery, invariant sequence. In curricular areas where this pattern is implemented, students master sets of objectives at their own rates. Within the units of instruction, the objectives may or may not be sequenced invariantly. In the Wisconsin Design for Reading Skills Development three of six areas (Word Attack Skills, Study Skills, and Comprehension Skills) incorporate Pattern 1 of instructional programming. Pattern 1 is also found in two other curricula developed at the Wisconsin Research and Development Center--pre-reading and mathematics. All five skill areas of the Pre-Reading Skills Program (PRS) and most of the units of instruction dealing with arithmetic in Developing Mathematical Processes (DMP) follow this pattern.

In the Wisconsin Design for Reading Skill Development objectives within each area are organized according to levels (e.g., four levels in the Word Attack area). Within each level the objectives are not sequenced invariantly. It is expected that each student will master all the objectives within a given area before proceeding to the next level; however, the student does not need to pursue the objectives within the level in an invariant sequence. A similar structure is found in the published science curriculum, Science: A Process Approach (SAPA). The present study deals with science concepts taught at Level 6, corresponding to the sixth year of schooling (assuming that Level 1 instruction begins in kindergarten). As with other curricula taught according to Pattern 1, students are expected to attain objectives within each unit of SAPA instruction to a specified criterion of mastery before going to the next unit of instruction; however, students do not encounter the units of instruction (or individual SAPA lessons) within a given level in an invariant sequence.

Pattern 2: Common objectives, full mastery, variable sequence. In this pattern instruction is organized so that all students pursue the same units of instruction to attain objectives to full mastery; however, the mastery of objectives for one unit is not prerequisite for starting another one. This means that a student may start any one of the planned units of instruction at any time during a semester or year. Students complete the units at their own rate; there-

fore, not all students complete the same number of units during the school year. This pattern is not a common one in IGE schools.

Pattern 3: Common objectives, variable attainment, invariant sequence. According to this pattern, all students pursuing a particular unit of instruction attain the same objectives to a criterion appropriate for each student. The units are sequenced invariantly during the school year. Some units of instruction in any curricular area follow this pattern in which learners may go to the next unit whether or not they have mastered the objectives of the prior unit.

Pattern 4: Common objectives, variable attainment, variable sequence. Individually Guided Motivation (Klausmeier, Jeter, Quilling, Frayer, & Allen, 1975), three areas of the Wisconsin Design for Reading Skill Development (Self-Directed, Interpretive, and Creative Reading Skills), and most of the geometry topics of Developing Mathematical Processes follow this pattern. All students are to attain the same objectives; they need not attain them to the same specified level, but to a level judged appropriate for each student; and students need not take the units in a fixed sequence.

Pattern 5: Variable objectives, full mastery, invariant sequence. Instruction may be arranged so that not all students attain the same objectives; however, sets of objectives in each unit of instruction taken by any single student are attained to mastery. Further, the units of instruction are undertaken in an invariant

sequence. An example of this pattern of instruction is one commonly found in instrumental music education. Here the students playing different instruments have different objectives; however, those learning to play the same instrument take individual units (e.g., lessons on sight reading, notation, scales) in an invariant sequence. Generally, students must reach the same level of mastery or proficiency before proceeding from one unit of instruction to the next.

Pattern 6: Variable objectives, full mastery, variable sequence. Instruction follows the same pattern here as in Pattern 5, except the units are not sequenced.

Pattern 7: Variable objectives, variable attainment, invariant sequence. This pattern is uncommon because of the invariant sequence. Although no particular level of attainment of objectives is required, the invariant sequence means that some knowledge or skill is needed before moving from one unit of instruction to the next.

Pattern 8: Variable objectives, variable attainment, variable sequence. Certain units of instruction in Developing Mathematical Processes which deal with probability and statistics follow this pattern. Pattern 8 of instructional programming is quite common; units of instruction are often "elective" in that not all students attending a school are expected to pursue these units, and those who do pursue the units are not required to reach any specified criter-

ion of attainment.

It can be seen that four of the eight patterns (1, 2, 5, and 6) of instructional programming outlined above include mastery as the criterion of attainment. Naturally, in order to develop mastery learning in students, one must be able to recognize when they have achieved it. "Teachers must be able to define what they mean by mastery, and they must be able to collect the necessary evidence to establish whether or not a student has attained it" (Bloom, Hastings, & Madaus, 1971, p. 52). Thus, monitoring the learning with formative tests, or diagnostic-progress tests as they are called, is a critical step in the IPM. In reviewing a large share of the mastery literature, Peterson (1972) reports that researchers have required anywhere from 75% to 100% mastery on the progress tests. One study by Airasian (1967) required 75% mastery on the formative tests; Kim (1970, 1971) required 80% mastery on assessments; the Keller (1968) approach typically required a perfect score, or 100% mastery, on unit tests. In the Wisconsin Design for Reading Skill Development 80% is used as the criterion of mastery in the areas of Word Attack and Study Skills. Although an 80% level of mastery is also applied in Developing Mathematical Processes, a second category called "Progress Towards Mastery" (60%-80%) is used for children who have not attained full mastery but are still allowed to move ahead to the next unit of instruction.

A key procedure in the mastery-learning approach is reteaching

or remediation for students who have not achieved mastery; this procedure is also incorporated in four of eight patterns of instructional programming for the individual student (Klausmeier, 1976, 1977). Following a diagnostic progress test, there are a number of resources available for remediation of non-mastery: small-group study sessions, tutorial help, alternative texts, workbooks, worksheets, academic games, audio-visual methods, and reteaching.

It is important to note that all features of the Bloom/Block Learning-For-Mastery approach are built into the Instructional Programming Model: formal specification of instructional objectives for each learning unit (Step 4 in the IPM), instruction (Step 5 in the IPM), formative/diagnostic evaluation after instruction (Step 6 in the IPM), and corrective or remedial instruction for students where needed (Step 7 in the IPM). Though all features of Learning-For-Mastery are to be found in the IPM, some aspects of instructional programming are totally lacking in mastery strategies such as those outlined by Bloom (1968, 1976), Block (1971, 1974), and Block and Anderson (1975). In particular, the IPM goes beyond the scope of the Bloom/Block conceptualization by specifying educational objectives for the student population of an entire building (Step 1), and by specifying a range of objectives that may be attainable for subgroups of the student population (Step 2); further, the IPM specifies pretesting each child to establish entering levels of achievement and motivation (Step 3). This preassessment is an important

part of instructional programming for the individual student. It allows either the placement of students into groups, or total individualization of instruction, relative to cognitive or affective entry characteristics. In IGE schools, evaluation of students' learning characteristics and achievements is aimed at providing information at three times: (1) at the beginning of a unit of instruction, (2) during the instructional sequence, and (3) at the end of a unit of instruction. Evaluation for decision making requires assessment at all three points in the educational process. As mentioned by Klausmeier:

> Evaluative information too often has been gathered to compare individuals and groups rather than to plan and carry out good instruction for individual students. As a result, it has been done after an instructional sequence is completed rather than before it is started and during the sequence. (1977, p. 18)

Only by pretesting can the individual needs of students be determined and taken into account during instruction. Within the framework of instructional programming for the individual student, instructional materials, use of space, allocation of time, mode of instruction (such as totally individual, or in small groups), type of instructional activity, and amount of teacher direction are all varied to meet the needs of the student.

COMPONENTS OF INSTRUCTIONAL PROGRAMMING IN THE PRESENT STUDY

Two basic problems associated with most mastery-learning strategies are finding a way to reduce the amount of time the slower

students require to a point where it is not prohibitively long, and also arranging new conditions in schools so that teachers can adapt rates and methods of instruction to the characteristics of individual students. Pattern 1 of instructional programming (Klausmeier, 1977) incorporates the same ideas of mastery as those of Carroll (1963). Further, the grouping of children in Instruction and Research Units of 75 to 150 children in IGE schools, rather than in age-graded classrooms, permits teachers to adapt both the amount of time for instruction and also the methods of instruction so that mastery learning in any curricular area should be possible. These three key ingredients of instructional programming for the individual student are included in the present study: (1) grouping for instruction on the basis of entry characteristics, e.g., achievement, motivation, learning style, (2) differential pacing, and (3) mastery learning according to Pattern 1 of instructional programming for the individual student. A brief review of research related to each of these three aspects of the IPM follows.

Grouping for Instruction

Inflexible, relatively permanent, homogeneous grouping on the basis of intelligence or general ability is being rejected by many educators (Ekstrom, 1961; Esposito, 1973; Heathers, 1969). Esposito defines homogeneous grouping as "the organization of instructional classes on the basis of students' similarities on one or more specific characteristics" (p. 165). He stresses that ability grouping

is not the only form of homogeneous grouping; another form is flexible grouping according to the skill needs of students in different learning activities. This approach was used by Willcutt (1969) to adapt instruction to individual needs in a junior high mathematics class. One-half of the seventh-grade students were assigned to classes on the basis of pretest scores for the mathematics unit to be taught: students in the lowest quartile formed one "review" class; students in the middle two quartiles formed two "standard" classes; and students in the top quartile formed one "depth" class. All classes received the same basic instruction, but the degree of emphasis differed in each group. Groupings were changed for each of eight units in mathematics, based on cognitive entry characteristics (i.e., pretest scores) for each unit. Teachers were allowed to arrange instruction as they customarily did. Control children were assigned to non-differentiated classes and taught in typical self-contained classrooms. Results on achievement posttests did not reveal any significant difference between experimental and control groups. However, an attitudinal survey revealed that significantly greater positive attitudes toward mathematics existed in the experimental groups. Willcutt (1969) suggested that a possible reason for the lack of significant differences in achievement may have been failure on the part of experimental teachers to effectively differentiate instruction according to the needs of the students.

An experiment was performed by Dewar (1963) to assess the

effects of intra-class grouping for mathematics instruction. Children in the experimental classrooms were grouped on the basis of previous standardized test scores and teacher judgment; three separate groups within each classroom were formed for instruction. Students in the matched control schools were not grouped for differentiated instruction. An alternate form of the Stanford Achievement Test in arithmetic was used as the dependent measure. An analysis of gain scores showed that the lowest and highest experimental groups made significantly greater gains that the lowest and highest control groups (control Ss were grouped only for the purpose of data analysis and not for instruction); there was no significant difference between experimental and control middle groups. Attitude surveys administered to the experimental students revealed positive feelings toward the grouping procedure.

Schilling (1976) points out that a major shortcoming in most studies of homogeneous grouping is that actual modifications of instruction to adapt to the specific needs of the groups are not made. She cites a study by Justman (1968) in which no attempt was made to adapt instruction to the needs of homogeneous groups; it was found, not surprisingly, that there was no relationship between the degree of homogeneity in a class and the amount of increase in reading achievement for that class.

Differential Pacing

An important component of instructional programming for the

individual student is the differentiation of instruction for groups of students with similar skill needs. In respect to pacing, Carroll (1963) states that aptitude is specific to the task under consideration. Aptitude may depend on, among other things, "the amount of prior learning which may be relevant to the task under consideration. A learner who has already progressed far toward the mastery of a task may not need much time to complete its learning" (p. 726). In Developing Mathematical Processes, the elementary mathematics program for Individually Guided Education, there is an acceptance of the fact that students achieve at different rates. Romberg (1977) describes the importance of pacing in IGE mathematics:

> Some students learn almost spontaneously, while others learn only after much repetition and belabored effort. The effective way to deal with different achievement rates is to pace instruction according to the differential rates. Mastery teaching of the mathematical processes is essential, but the pace of such teaching must vary greatly from individual to individual. (p. 78)

Washburne (1932) of the Winnetka School System stated that time has been the one measurable factor in education. The educational system is laid out in terms of time units. Traditionally, these time units have been held constant, while the achievement of the students has varied according to individual abilities. Carroll suggested that educators make units of achievement the constant factor, while varying the time to fit the individual capabilities of the learners.

Pretest scores may be considered a valid indication of the

amount of time needed for mastery. Yeager and Kissell (1969) investigated the relationship between a composite of student entry characteristics and time needed to complete a unit in mathematics for nine- and ten-year-old children in an elementary school. The hypothesis was that there exists a significant positive relationship between a student's initial state before beginning to work on a given task and the amount of time required to complete the task. The study examined the intercorrelations between the criterion (number of days needed to master the objectives of the unit) and five variables in student entry characteristics. The variables were: (1) unit pretest scores; (2) number of skills to be mastered in the unit; (3) student's IQ score; (4) student's chronological age; and (5) total units previously mastered. Results indicated that the number of days needed to master a unit was significantly related to the student's entering state, with the most important factor being the student's unit pretest score. The next two most important factors were the number of skills to master in the unit and the student's age. The least important variable was IQ.

The amount of time needed for mastery may vary widely for different individuals. Glaser, Reynolds, and Fullick (1963) studied the use of programmed instructional materials within the structure of an ongoing classroom situation. A program in recognizing numerals was conducted with 121 first-grade children. It was found that the amount of time needed for completion of the program varied from

8 to 25 days. When time needed for completion was compared to post-test scores, little correlation was found between the two variables. Most children achieved mastery in the program, indicating that with sufficient time available, children could reach mastery on the learning task.

In another study with young children, Suppes (1964) investigated the time required to master beginning reading tasks. Kindergarten children were trained to criterion (28 successively correct responses) on a word recognition task. After this, 14 two-word phrases, then three-word, four-word, and five-word phrases were learned by the children to criterion. Thirty-eight children completed the experiment. In reporting the results, Suppes called attention to the wide range in total number of trials required: the fastest child needed only 196 learning trials, while the slowest child required 2,506. (The range was not misleading, for the mean was 967.4, with a standard deviation of greater than 400.) This study illustrates the great difference in time needed to attain mastery of some learning tasks. Suppes concluded that concentration on individual differences in time needed for learning would be necessary for the improvement of subject matter learning.

Mastery Learning in IGE (Pattern 1)

As discussed earlier, Pattern 1 of instructional programming is one in which mastery of the same objectives is to be attained by all students and in which the units of instruction are to be taken in an

invariant sequence. This pattern is to be found in two programs developed by the Wisconsin Research and Development Center for Cognitive Learning and implemented in numerous schools throughout the United States. These two curricular programs are the Wisconsin Design for Reading Skill Development (Design) and Developing Mathematical Processes (DMP).

The Design is an objective-based system that provides both structure and substance for an elementary school reading program, kindergarten through grade six (Otto, 1977). It has been shown to be an effective program in schools that are diverse in achievement levels, organizational plans (e.g., IGE or traditional), and geographic locations (Askov, 1970; Chester, Hubbard, & Van Blaricom, 1975; Johnson, 1971; Klumb, 1973; Quilling, 1971; Quilling & Wojtal, 1972; Rude, 1973). Student achievement scores on the Design tests as well as standardized tests (administered before and after instruction with Design materials) have favored users (Askov, 1970; Chester et al., 1975; Quilling & Otto, 1971; Rude, 1972). As described earlier, one area of the Design which is taught according to Pattern 1 of instructional programming for the individual student (Klausmeier, 1977) is Study Skills. A report (Sals & Hubbard, 1975) on the results of the Study Skills field test indicates that such a pattern of instruction is an effective approach. Consistent improvement was shown by Study Skills students on program-embedded and standardized measures. In all analyses, schools with lower typical

reading achievement showed greater improvement than schools with higher typical reading achievement. In five schools where reading scores were typically lower, the students using Study Skills materials according to Pattern 1 of instructional programming were reported to have exceeded non-participating students on 95% of the program objectives; mean difference in grade equivalents between baseline and follow-up scores was six-tenths of a grade higher for participating students than for non-participating students. In seventeen schools where reading scores were typically high, those students using the Design according to Pattern 1 were reported to have exceeded non-participating students on 76% of the program objectives; mean difference in grade equivalents between baseline and follow-up scores was one-tenth of a grade higher for participating students than for non-participating students.

Another curricular program which incorporates Pattern 1 of instructional programming for the individual student is Developing Mathematical Processes. An evaluation study (Wearne & Whitaker, 1976) was conducted near the end of the 1975-76 school year by developers of DMP to assess the performance of students who had been using the program materials. Standardized tests were administered in ten states to children in 25 schools. Results were as follows: the percentile rank of the mean score for grade 2 children on the Cooperative Primary Test was 65, and the percentile rank of the mean score on the California Achievement Test was 51; for grade 3 chil-

dren who took the Iowa Test of Basic Skills, the percentile rank of the mean score was 63. The percentile rank describes the performance of DMP children relative to all children taking the test nationally.

Because certain items in the standardized instruments involve mathematical concepts not formally included in DMP objectives, a further analysis of test scores was performed on subsets of items (Wearne & Whitaker, 1976). Two subsets of items were selected for consideration: (1) items assessing behaviors which most children could be expected to master in DMP topics completed during the year, and (2) items assessing behaviors which were introduced or alluded to in a DMP topic, but for which mastery was not expected in the program. The first set of items was designated the "M" subset (for "Mastery") and the second set was the "P" subset (for "Progress Towards Mastery").

When the second-grade data were analyzed according to subsets, it was found that grade 2 children in the DMP program answered 72% of all "M" items correctly on the Cooperative Primary Test; of the "P" items, second-graders answered 62% correctly on the CPT. On the California Achievement Test, second-graders in the program correctly answered 88% of all "M" items and 90% of all "P" items. Third-grade students receiving DMP instruction correctly answered 69% of all "M" items and 60% of all "P" items on the Iowa Test of Basic Skills.

In summarizing the effectiveness of the DMP program, Moser and

Romberg (1977) state:

> We do not claim it is the mathematics program for all teachers and children--, but we know DMP works, its mathematics is sound, educators are impressed, children like it, and there is lots of evidence that children learn the basic concepts and skills of mathematics. (p. 13)

Evidence from two comprehensive programs incorporating Pattern 1 of instructional programming--the Design and DMP--suggests that in curricular areas where this pattern (common objectives, full mastery, invariant sequence) has been implemented, the achievement of students is consistently high.

III

METHOD

The present study was conducted within the framework of a larger, longitudinal intervention study of children's learning in science. The purpose of the study was to examine the effects of a method of instruction based on instructional programming for the individual student. In the experimental treatment, students were grouped for differentiated instruction on the basis of entry levels of achievement, and time and instructional techniques were varied depending upon the needs of the independent groups.

DESCRIPTION OF SAMPLE

Schools

The two experimental and two control schools involved in the intervention study served as experimental and control schools in the present study. These schools were originally selected in conjunction with administrators of the participating school system. A total of four elementary schools were selected; all schools were implementing IGE and were located in the same small, midwestern, industrial city. Two schools with nongraded classrooms were selected: one (School E_1) randomly assigned to the experimental condition, and the other (School C_1) to the control condition. A pair of schools with graded classrooms

was also selected: one (School E_2) randomly assigned to the experimental condition, and the other (School C_2) to the control condition. All four schools were matched as nearly as possible on those socioeconomic, ethnic, and general demographic characteristics representative of the city of Janesville, Wisconsin. Fifty to eighty percent of the school populations fell into the middle and higher-than-middle socioeconomic strata. Blue collar, business, and professional occupational categories were represented by the parents of the children attending the schools. In each of the four IGE schools, there were from 60 to 80 students at level six (equivalent to the sixth year of schooling) and from one to three teachers of science at level six.

Subjects

The subjects in the present study were drawn from approximately 265 students at level six attending the four elementary schools. A total of 235 children were selected as subjects. Excluded were students who had been absent for all or part of the pretesting at the beginning of the semester, or who had missed more than one of the six subtests given during posttesting. Three of the 235 subjects missed only one subtest during posttesting and were retained for data analysis; a mean for the missed subtest was substituted for the student's score, based on the performance of all other students in the same group within a particular school on the given subtest. Of the 235 subjects selected, approximately 15 missed one of the diagnostic-

progress tests given at the end of each unit of instruction throughout the semester; however, these subjects had been present for all pre- and posttesting, and were retained for data analysis. While 235 students were selected as subjects, the other students in the four schools also participated fully in the instructional treatments and were excluded only for the purpose of data analysis.

In the present study subjects were assigned to groups (high, middle, or low) on the basis of entry levels of achievement. Scores on the 60-item pretest covering the six units of SAPA science were used to divide students in each experimental school into three classroom-size groups for instruction. The rationale for determining cut-off scores for the high, middle, and low groups was a pragmatic one--instructional groups had to be of manageable size. Cut-off points in the distribution of pretest scores for experimental subjects were as follows: students with raw scores of 35 and above were assigned to the high group; students with raw scores in the 27-34 range were assigned to the middle group; and students with raw scores of 26 and below were assigned to the low group. In the experimental schools, the high, middle, and low groups were taught in separate classrooms by different teachers, and students received differentiated instruction according to the needs of the group. The sizes of the instructional groups within each experimental school were roughly the same. As shown in Table 1, at School E_1 there were 20, 20, and 20 subjects in the high, middle, and low groups, respectively; at School E_2 there

TABLE 1

NUMBER OF SUBJECTS IN EACH GROUP AT EACH

OF THE FOUR PARTICIPATING SCHOOLS

Group	School E_1	School C_1	School E_2	School C_2	Total for Experimental Schools E_1E_2	Total for Control Schools C_1C_2
High (Pretest scores 35 and above)	20	13	24	18	44	31
Middle (Pretest scores 27 - 34)	20	16	18	24	38	40
Low (Pretest scores 26 and below)	20	27	23	12	43	39
Total Across Groups	60	56	65	54	125	110

were 24, 18, and 23 subjects in the high, middle, and low groups, respectively.

While pretest scores were used to organize groups for instruction in the experimental schools, they were not used in the same manner for control subjects. In the control schools the level of pretest score (high, middle, or low) was used only for the purpose of data analysis and not for organizing instruction. The same raw score cut-offs were used for the control subjects as for the experimental subjects. It can be seen from Table 1 that this resulted unavoidably in an unequal distribution of subjects among the three groups in the control schools: at School C_1 there were 13, 16, and 27 subjects in the high, middle, and low groups, respectively; at School C_2 there were 18, 24, and 12 subjects in the three groups, respectively.

Of the total 235 subjects in the study, there were 125 in the experimental condition and 110 in the control condition. A breakdown of the participating Ss into groups by sex has not been performed inasmuch as sex differences have not been reported in previous studies (Klausmeier, Allen, Sipple, & White, 1976; Klausmeier, Sipple, & Allen, 1974).

INSTRUCTIONAL MATERIALS

The elementary science curriculum used in the participating school system is SAPA (Science: A Process Approach). This "process" approach to science education was developed by the American Association for the Advancement of Science (1967), and later marketed by

Xerox Corporation. Although a great deal of "content" is stressed in individual units of instruction, the broad aim of SAPA is to promote understanding of the processes of science (i.e., observing, classifying, inferring, predicting, measuring, using numbers, using space/time relationships, communicating, formulating hypotheses, interpreting data, experimenting). In the longitudinal intervention study, printed lessons were developed to teach these process-concepts of science. There are no student textbooks and no printed lessons for the SAPA-content units of instruction. A teacher's guide for each lesson specifies behavioral objectives and suggests instructional activities to be pursued. The classroom teacher is responsible for directing activities, performing demonstrations, supervising experiments, asking questions, and leading discussions.

SAPA Lessons

The present study involved six of the eight SAPA lessons normally taught at level 6 during second semester (presented in Table 2). Instruction on the two lessons in the Controlling Variables block had been completed prior to the onset of the present experiment. As shown in Table 2, the six SAPA-content lessons used in the present study were lessons L, M, N, O, P, and Q. Reference letters used to designate the lessons are not associated with the published curriculum, but were adopted in the longitudinal intervention study for easy identification.

TABLE 2

SAPA SCIENCE LESSONS TAUGHT AT LEVEL SIX DURING SECOND SEMESTER OF INSTRUCTION

Reference Letter	SAPA Lesson Title	SAPA Lesson Classification	Part	Booklet
J	Rolling Cylinders	Controlling Variables 1	E	b
K	Upward Movement of Liquids	Controlling Variables 2	E	c
L*	Guinea Pigs in a Maze	Interpreting Data 1	E	d
M*	Observations & Hypotheses	Formulating Hypotheses 1	E	h
N*	Suffocating Candle	Predicting 4	D	m
O*	Using Punch Cards	Classifying 10	D	t
P*	Describing Location	Communicating 11	D	j
Q*	Reporting an Investigation in Writing	Communicating 12	D	r

*Lessons used in the present study.

Lesson L, or Interpreting Data 1, is called "Guinea Pigs in a Maze." In this unit of instruction, children collect and interpret data related to the behavior of two guinea pigs in a maze. Behavioral objectives for this lesson are as follows: the child will be able to (1) describe the changes in an animal's performance of a task which result from repeated trials or practice; (2) construct a graph showing how an animal's performance depends on the number of trials or the amount of practice; (3) identify the data that support the statements that the time required to perform an act will, in general, be shortened after a number of trials; and (4) construct answers to questions about information presented in a graph.

Lesson M, or Formulating Hypotheses 1, is titled "Observations and Hypotheses." In the SAPA curriculum, an hypothesis is defined either as an inference that is generalized to include all objects or events of the same class, or as a generalization about observations. In this exercise, children make simple observations to suggest some broad generalization (i.e., hypothesis). They then examine further observations to determine whether or not these observations support the hypothesis. Behavioral objectives for this lesson are as follows: the child will be able to (1) distinguish between statements that are hypotheses and those that are not; (2) distinguish between observations that support an hypothesis and those that do not; and (3) construct an hypothesis from a given set of observations.

Lesson N, or Predicting 4, is called "The Suffocating Candle."

In this unit of instruction, the children first observe a burning candle that is covered by a glass jar. Then they predict whether the burning time will be longer or shorter with different jars, making predictions based on measurements of both the jar volumes and burning times. In addition, the students are introduced to the techniques of interpolation and extrapolation from data in a graph. The stated behavioral objectives are: the child will be able to (1) construct predictions based on a series of observations that reveal a pattern; and (2) construct a revision of a prediction on the basis of additional information.

Lesson O, or Classifying 10, is titled "Using Punch Cards to Record a Classification." This exercise introduces a manual punch-card system that is useful in recording simple classifications or yes/no data. Children learn to notch (or cut a "V") when the answer to a question is "yes" and to sort cards with a large needle. This lesson involves a great deal of individual and group participation; once students gain skill in using the classification system, they are encouraged to make up their own classifications and to sort decks of cards to tabulate categorical data. Behavioral objectives for this lesson are as follows: the child will be able to (1) construct a classification of objects by using at least one series of subsets; (2) construct a punch-card system to record simple data; and (3) demonstrate the use of a punch-card system.

Lesson P, or Communicating 11, is called "Describing Location."

This lesson provides instruction in the use of a Cartesian coordinate system--the graphing of points on a grid with intersecting x and y axes. Prior exposure to the concepts in this SAPA lesson was provided in those topics of the math curriculum related to graphing. Behavioral objectives are: the child will be able to (1) name the ordered number pair that locates a position on a graph or a grid; and (2) identify the position of a number pair on a graph or a grid.

Lesson Q, or Communicating 12, is called "Reporting an Investigation in Writing." The purpose of this lesson is to introduce the students to reporting the procedures and the results of an investigation. The content of this lesson is highly non-specific, and the skills taught in this unit are more akin to language arts than to science. Nevertheless, it is important for the scientist to be skillful in communicating his or her ideas and to share results with others in the scholarly community. At the end of this unit of instruction, the child should be able to describe in writing all parts of an investigation that (s)he has just observed or conducted, including the purpose, method, materials, procedures, and results.

SAPA-Content Tests

In the present study the dependent measure used to assess knowledge of SAPA content was a 90-item multiple-choice, objective test; the SAPA-content test consisted of six 15-item subtests (L, M, N, O, P, and Q). Ten-item versions of these subtests had already been developed and validated for use in the longitudinal intervention study.

The experimenter lengthened each subtest by 50% in order to improve test reliabilities. As shown in Table 3, Hoyt internal consistency reliabilities for the six expanded subtests ranged from .72 to .77; when calculated for the entire set of 90 items, the reliability coefficient was .93. The six SAPA-content subtests are presented in Appendix A.

INSTRUCTIONAL PROCEDURES

Experimental Schools

In the two experimental schools, the treatment condition was a method of instruction based on instructional programming for the individual student (Pattern 1). In this method the attempt was made to secure mastery of the same science objectives by as many students as possible within the time constraints of the experiment. Instead of grouping children in heterogeneous classrooms for instruction, teachers arranged instruction for relatively homogeneous groups (high, middle, and low, based on pretest level of achievement), where both the amount of time for instruction and the instructional tactics and materials were varied for the homogeneous groups. Due to the fact that science instruction for all six units had to be completed by the end of the semester, it was not anticipated that all the students in the low group would have sufficient time to attain mastery of all lesson objectives by the end of the experiment.

As detailed earlier in this chapter, each experimental subject was assigned to a high, middle, or low instructional group, based on

TABLE 3

PSYCHOMETRIC CHARACTERISTICS OF THE SIX SAPA-CONTENT TESTS

ADMINISTERED TO EXPERIMENTAL AND CONTROL GROUPS (POSTTEST)

SAPA-Content Test	Number of Items	Mean	Mean as %	Standard Deviation	Hoyt Internal Consistency Reliability	Standard Error of Measurement
L. Guinea Pigs in a Maze *(Interpreting Data 1)*	15	11.62	77.45	2.71	.74	1.32
M. Observations & Hypotheses *(Formulating Hypotheses 1)*	15	10.62	70.78	3.21	.77	1.51
N. Suffocating Candle *(Predicting 4)*	15	10.35	69.02	3.11	.74	1.52
O. Using Punch Cards *(Classifying 10)*	15	11.63	77.55	2.66	.72	1.35
P. Describing Location *(Communicating 11)*	15	11.72	78.10	2.65	.73	1.33
Q. Reporting an Investigation *(Communicating 12)*	15	10.15	67.69	2.99	.77	1.42
Composite of all tests L-Q	90	66.04	73.38	14.18	.93	3.62

pretest level of achievement. On the 60-item pretest, children with raw scores of 35 and above (high group) were assigned to a single teacher for the 10-week duration of the study, as was the case for subjects in the 27-34 raw score range (middle group) and subjects with raw scores of 26 and below (low group). Within each experimental school, teachers were randomly assigned to instructional groups. There were three teachers at level 6 in each of schools E_1 and E_2.

For each instructional group within an experimental school the major objective was to move through the SAPA lessons at a pace that gave most students the opportunity to master the material. Elements of instruction for each SAPA lesson included: (1) initial teaching of the content; (2) administration of a diagnostic-progress test to determine students' levels of mastery; (3) immediate feedback and review of test results; and (4) reteaching in small or large groups for students needing additional instruction. Table 4 summarizes the elements of instruction that were present in experimental and control schools. A discussion of the four elements as applied to each unit of instruction in the experimental classrooms is provided below.

1. Initial teaching. All teachers followed the teacher guides that outlined activities to be undertaken and concepts to be presented for each SAPA lesson. In every experimental classroom, the teacher used a vocabulary chart to teach the key vocabulary words. The length of initial instruction was determined by the teacher, based on individual judgment.

TABLE 4

SUMMARY OF INSTRUCTIONAL CONDITIONS FOR EXPERIMENTAL AND CONTROL TREATMENTS

Treatment	60-item SAPA-content Pretest	Grouping Based on Pretest	Initial Instruction	15-item Diagnostic-Progress Test During Semester	Review of Test and Reteaching of Concepts Not Mastered	90-item SAPA-Content Posttest
Experimental	X	X	X	X	X	X
Control	X		X	X		X

When a teacher judged four-fifths of the students in the group to have mastered most of the concepts in the lesson, the initial teaching phase was concluded and a diagnostic-progress test administered.

2. Diagnostic testing. All students in the group took a 15-item SAPA-content subtest corresponding to the SAPA lesson being taught. Mastery was defined as 80% of the 15 items answered correctly. Students marked their item responses on answer sheets which were given to an instructional aide for grading.

3. Feedback and review. After the diagnostic-progress tests were graded, the teacher returned the tests to the students and every item that was missed by two or more students was discussed.

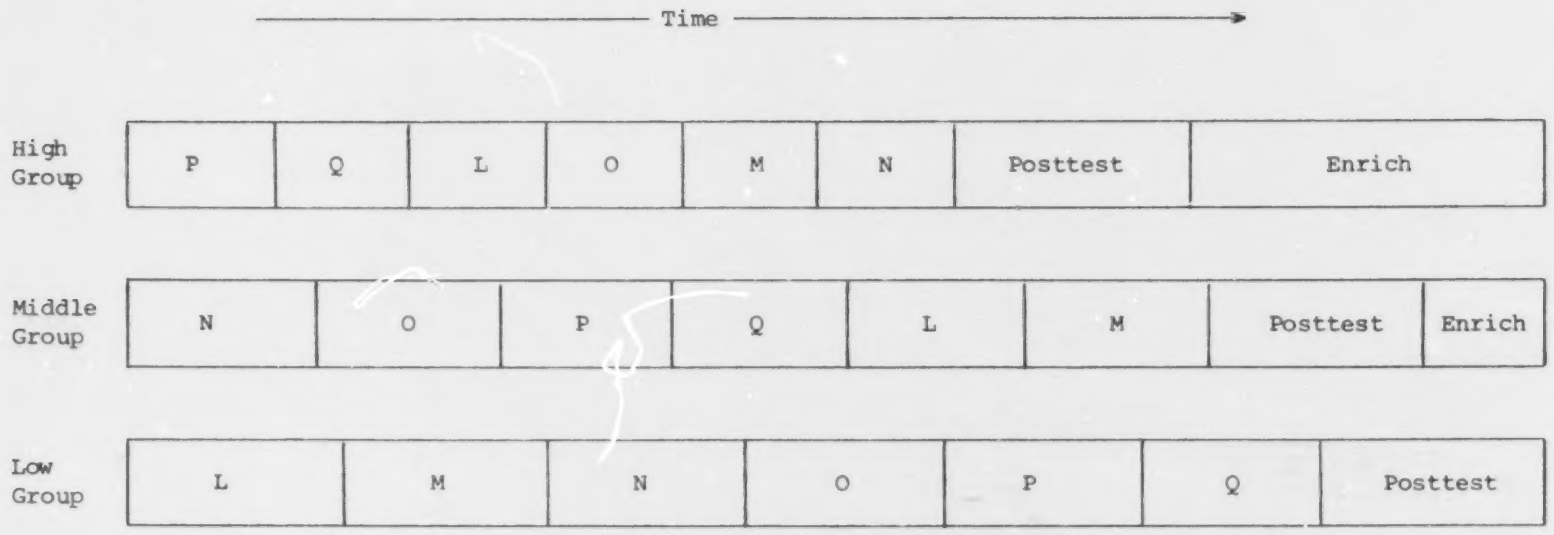
4. Reteaching. Following feedback to the students about their performance on the test, the teacher spent from 15 to 60 additional minutes reteaching concepts that were not learned. Students who had demonstrated mastery on the diagnostic test sometimes tutored individually or in small groups; in some cases reteaching was accomplished in large-group discussion or by providing worksheets related to those concepts most frequently missed.

As soon as reteaching was completed, the group moved immediately to the next SAPA lesson. For each of the six SAPA lessons taught during

the study, the complete cycle of initial teaching, diagnostic testing, feedback, and reteaching was accomplished.

A pragmatic concern in both experimental schools was the availability of apparatus and materials for each SAPA lesson. In order to ensure adequate equipment and supplies, the six lessons were sequenced differently for each group within a particular school; this enabled the staff to have access to the needed materials without exhausting the limited resources already available. The sequence of SAPA lesson presentation for high, middle, and low groups within each experimental school is shown in Figure 4.

In the experimental schools, materials and instructional techniques were tailored to the needs of the instructional groups; for example, quick progress through the basic SAPA lesson may have been appropriate for the high group, while more detailed explanation with supplementary texts or visual aids could have been used for the low students. Enrichment activities (other than simple reteaching and review of the basic content) were postponed until the completion of all SAPA lessons and final posttesting. The method of instruction in experimental schools provided for flexibility in the pacing of units of instruction. Each group was allowed to move through the sequence of lessons independently of other groups in the school, with the high group expected to spend the least amount of time and the low group the greatest amount of time. Teachers were permitted to employ whatever teaching techniques they felt were appropriate to



Lesson Key:

L -- Interpreting Data 1 (Guinea Pigs in a Maze)
M -- Formulating Hypotheses 1 (Observations & Hypotheses)
N -- Predicting 4 (Suffocating Candles)
O -- Classifying 10 (Using Punch Cards)
P -- Communicating 11 (Describing Location)
Q -- Communicating 12 (Reporting an Investigation)

Figure 4. Sequence of SAPA lesson presentation for high, middle, and low groups in experimental schools.

effective instruction (i.e., small-group work, work in pairs, teacher demonstrations, audio-visual methods, tutoring, etc.).

Summary of Experimental Treatment. Six elements of the Instructional Programming Model (IPM) as they were implemented in the present study are outlined below:

Step 1. Educational objectives were specified for the student population of a building. In the present study, the school-wide objectives for science were the objectives of the SAPA lessons specified for kindergarten through sixth grade.

Step 2. The objectives that may be attainable for subgroups of the student population were determined. The subgroup of the school population selected for the present study was level six; the objectives for science specified for the subgroup were the objectives of the SAPA lessons to be taught at level six.

Step 3. An assessment was made of each child's level of achievement, learning style, and/or motivational level. Each student's entering level of achievement was assessed by a pretest at the beginning of the second semester. The 60-item pretest measured each child's pre-instructional knowledge related to the six SAPA lessons L, M, N, O, P, and Q. In the present study, learning style and motivational level were not assessed.

Step 4. Instructional objectives were set for each child for a specified period of time. In this study the same objectives were set for every child for the period of time during which the lesson was

taught. The period of time was variable from group to group within the school.

Step 5. An instructional program suitable for each student was planned and implemented. For each SAPA lesson, the instructional procedures included (1) initial instruction, (2) diagnostic-progress testing, (3) feedback and review, and (4) reteaching concepts that were not learned. An attempt was made to bring as many students to mastery as possible within each group.

Step 6. Final assessment was made for attainment of objectives. In the present study a final posttest consisting of 90 items was administered to all children at the conclusion of SAPA instruction. No enrichment or related activities were provided for any instructional group until after the final posttest had been administered.

The writer assumed sole responsibility for teacher orientation and monitoring of conditions of instruction throughout the study. Early in February, 1977, separate meetings were conducted for the experimental and control teachers. During the meeting for the six experimental teachers, sealed envelopes were distributed randomly to the E_1 and E_2 staffs, indicating assignment to one of three groups (high, middle, or low). Each teacher learned which group (s)he would teach and received an appropriate roster of names based on a prior analysis of students' pretest scores. The meeting for E_1 and E_2 teachers consisted of an oral presentation by the writer and a question/answer period. The basic procedures to be followed in each experimental school were summarized in an outline distributed to each

teacher. (The instructions to experimental teachers are presented in Appendix B.)

Control Schools

In the two control schools instruction was given on the same SAPA units of instruction; however, the students were heterogeneously grouped in their regular science classes. Children with high, middle, and low entry levels of achievement were present in each class. Pretest performance was not taken into account when organizing for instruction in the control schools. As shown in Figure 4, control subjects received pretests, initial instruction, diagnostic-progress tests, and posttests; there was no feedback concerning test results and no reteaching. Time spent on instruction was roughly equivalent for each of the heterogeneously grouped classes. Control teachers presented the lesson content according to the SAPA teacher guides, and instruction continued until most children in the class were judged to have mastered most objectives. At the end of each unit of instruction, control students took a 15-item diagnostic test on the content of the SAPA lesson just completed. Immediately following the test, instruction on the next SAPA lesson commenced.

In School C_1 there were three teachers of level six science. Instruction was organized according to a technique called "round robin," whereby the six lessons were divided into two "rounds"; the first round consisted of lessons L, M, and N, and the second round consisted of lessons O, P, and Q. During one round of instruction, each teacher took responsibility for a single lesson and taught that

lesson three times successively to three different groups of children. At the end of round 1, for example, every child had received instruction on lessons L, M, and N; each teacher had provided instruction on one SAPA lesson to three different groups of children.

In School C_2 there was only one teacher of science at level six. The same teacher taught all three heterogeneous classes. The sequence of SAPA lessons at School C_2 was as follows: L, M, N, Q, O, P. (See Appendix B for directions to control school teachers.)

DATA COLLECTION

Three kinds of assessment instruments were used in the present study: (1) multiple-choice SAPA-content tests, to measure science achievement; (2) a Likert-type attitude inventory, to measure affective consequences of instruction; and (3) classroom observations and teacher reports, to document the ways in which instruction was carried out. A discussion of each type of data and method of collection is provided below.

SAPA-content Tests

In all four participating schools, SAPA-content tests were administered at three points: (1) at the beginning of second semester the six 10-item subtests were combined and given as a pretest of science achievement; (2) during the semester the 15-item versions were given as diagnostic-progress tests (subtests were administered separately following instruction on individual SAPA lessons with the sequence of administration corresponding to the sequence of teaching);

and (3) at the conclusion of all SAPA instruction the six 15-item subtests were combined and given as a posttest of science achievement. In all schools, whenever SAPA-content tests were administered, the test questions were read aloud to the students by the classroom teacher.

As outlined above, the 90-item SAPA-content posttest was administered when a group of children had completed all six SAPA lessons. At School E_1 the high group finished instruction in eight weeks and took the posttest before the middle and low groups had finished instruction. The middle and low children at School E_1 completed instruction during the tenth week and took the posttest during the eleventh week. At School E_2 the high and middle groups completed instruction in ten weeks and were posttested during the eleventh week; the low group required an additional week of instruction to finish the last SAPA lesson and, therefore, was posttested a week later than the other two groups. Schools C_1 and C_2 completed all SAPA instruction in 10 weeks and administered the final posttest at about the same time as the middle and low groups in School E_1 and the high and middle groups at School E_2.

While pretest and posttest data were subjected to statistical analysis, SAPA-content data collected during the semester was not analyzed because time of administration, sequence of administration, and conditions of testing were so varied. Evidence from the longitudinal intervention data (Klausmeier, Sipple, Swanson, & Schilling,

1976) suggests that there is little difference between immediate testing after instruction and final posttesting at the end of a semester. In the present study, control Ss took the diagnostic tests after initial instruction to allow equal exposure to the test materials for subjects in the experimental and control conditions.

Attitude Survey

In order to assess the affective consequences of the experimental treatment, a short, self-report Science Attitude Inventory (SAI) was administered. This 24-item instrument (presented in Appendix C) was developed for use in the longitudinal intervention study and described in the first-year intervention report (Klausmeier, DiLuzio, & Brumet, 1976). The SAI was administered in December 1976 as part of first semester posttesting for the intervention study. The results of this administration of the SAI were used as a baseline for attitude scores in the present study. The SAI was administered again as a posttest in May 1977 just prior to the administration of the SAPA-content posttests.

As originally devised, only 16 of the 24 items were to be scored. The eight items omitted from scoring were numbers 2, 8, 10, 12, 14, 19, 21, and 23 (see Appendix C); these items were related to non-science attitudes, such as "I enjoy math better than spelling," or "I would like to learn to speak a foreign language well." In the present study, scoring of the attitude inventories was modified by eliminating four additional items. Analyses conducted on the May

posttest data revealed that items 4, 9, 16, and 18 had point biserial correlation coefficients of .25, .37, .41, and .46, respectively. (The biserial correlation coefficient is an indication of an item's discriminating power; the higher the coefficient, the greater the discriminating ability of the item.) In order to improve the internal consistency of the Science Attitude Inventory, all items with coefficients of .50 or less were eliminated from scoring. Test characteristics for the scored SAI items are summarized in Table 5. As shown in the table, the reliability coefficients of the 12 scored items ranged from .51 to .80. Cronbach's (1951) coefficient alpha computed for the 12 scored items was .91. All positive statements (items 3, 5, 7, 13) were weighted as follows: Agree = 4; Slightly Agree = 3; Slightly Disagree = 2; Disagree = 1. All negative statements (items 1, 6, 11, 15, 17, 20, 22, 24) were weighted in a reverse fashion: Agree = 1, Slightly Agree = 2; Slightly Disagree = 3; Disagree = 4. The highest positive score attainable was 48; the lowest possible score was 12.

Classroom Observations and Teacher Reports

All teachers participating in the study were asked to keep detailed notebooks, or "logs," documenting conditions of instruction for each unit of SAPA science. Every teacher in Schools E_1 and E_2 recorded the SAPA activities that were taught and the total time spent on instruction for each science period. They also recorded the type of activities (e.g., initial instruction, diagnostic testing,

TABLE 5

PSYCHOMETRIC CHARACTERISTICS OF THE SCIENCE ATTITUDE INVENTORY (POSTTEST)*

Item Number	Mean	Standard Deviation	Estimate of Reliability
1	3.44	0.82	.70
3	3.24	0.90	.76
5	2.77	1.04	.70
6	3.09	1.05	.65
7	2.06	0.97	.51
11	3.38	0.86	.65
13	3.14	0.93	.80
15	3.26	0.91	.79
17	.31	0.85	.80
20	2.26	1.09	.63
22	3.21	0.98	.79
24	2.90	1.14	.73
Composite (12 items)	36.06	8.15	.91

*Only 12 of the 24 items were scored.

review, enrichment), the methods of instruction (e.g., large-group, small-group, teacher demonstration, work in pairs), and the number of children present for instruction. In Schools C_1 and C_2 the teachers recorded the number of minutes spent on each unit of instruction, the number of children present, and the SAPA activities that were taught. All teachers were encouraged to record critical comments, professional observations, or special problems, as well.

In addition to the teacher logs, documentation of classroom instruction was obtained through observations by the Experimenter. Instruction and testing in each of the four schools were carefully monitored at least once per week for the duration of the study. This was done to supplement information provided in the teacher logs and to describe how instruction actually took place.

Reports of Time Spent on Instruction. One of the items reported on all teacher logs was the amount of time spent on instruction for each SAPA unit of instruction. Table 6 presents the number of minutes spent on instruction by each group within the four participating schools. It can be seen that instructional time varied widely among groups and from lesson to lesson for the two experimental schools. In terms of the average number of minutes spent on a unit of instruction, the experimental groups at Schools E_1 and E_2 were in the expected direction: high < middle < low. At School E_1 the mean number of minutes spent by the high, middle, and low groups was 97.5, 175.8, and 219.2, respectively. At School E_2 the mean number of minutes for

TABLE 6

SUMMARY OF TIME SPENT ON INSTRUCTION IN EACH EXPERIMENTAL AND CONTROL CLASSROOM

School	Instructional Group	Time for Instruction (in minutes) Per Unit						Mean Number of Minutes Per Unit
		L	M	N	O	P	Q	
	High	90	75	120	80	120	100	97.5
E_1	Middle	120	90	270	255	200	120	175.8
	Low	360	225	180	180	180	190	219.2
	Hetero-geneous	150	180	170	150	170	180	166.7
C_1	Hetero-geneous	150	180	180	150	170	180	168.3
	Hetero-geneous	150	180	170	150	170	180	167.5
	High	150	90	205	150	150	145	148.3
E_2	Middle	190	130	145	160	180	100	150.8
	Low	250	240	240	210	300	300	256.7
	Hetero-geneous	180	180	180	180	180	180	180.0
C_2	Hetero-geneous	180	180	180	180	180	180	180.0
	Hetero-geneous	180	180	180	180	180	180	180.0

the high, middle, and low groups was 148.3, 150.8, and 256.7, respectively.

In each control school the average number of minutes spent on each lesson was roughly the same as the mean for middle groups in experimental schools. At School C_1 a round-robin instructional technique was used (lessons L, M, and N were taught in one round and lessons O, P, and Q were taught in the second round). Since all three teachers at School C_1 had to rotate groups at the same time, it was suggested that as nearly the same number of minutes as possible be spent on each lesson in the round. Ideally, initial instruction would have taken approximately 180 minutes per lesson, as reported at School C_2. In School C_1, however, it happened that less time was taken for initial instruction on some units; notably, in the first round, the teacher of Lesson L took only 150 minutes while the teacher of Lesson M spent a full 180 minutes. Again in the second round, the teacher of Lesson O spent 30 minutes less on initial instruction than the teacher of Lesson Q. It was learned by the Experimenter that at School C_1, when one group finished initial instruction early the extra time was spent in reviewing for the diagnostic test; in some cases as much as 30 minutes were spent in review prior to testing (an instructional procedure that was not recommended for control schools in the present study). Teachers in School C_1 did not teach to the test; however, to the extent that instructional time was spent in review just prior to testing, instructional conditions in the two

control schools were not comparable.

Descriptions of Classroom Activities. Provided below are descriptions of classroom activities, based on entries in the teacher logs and direct observation.

School E_1, High Group. The teacher of this group moved quickly through the sequence of SAPA lessons, completing instruction in just eight weeks. There was limited teacher demonstration and little individual work, but extensive class discussion and experimentation. Typically, about 20-30 minutes at the beginning of each SAPA lesson were spent in vocabulary work and introduction by the teacher. Then students were organized in smaller groups of four to five for conducting investigations. Following administration of a diagnostic test, the teacher reviewed all items that were missed by two or more students. For five of the six SAPA lessons, less than 15 minutes were spent in reteaching after feedback and review of the test; for Lesson O approximately 45 minutes were spent in reteaching because of the relatively poor performance evidenced on the diagnostic test. In his log, the teacher noted that for Lesson P "Many kids in this group already had knowledge of integers and coordinate points through math classes." For Lesson Q he noted: "These kids are skilled in use of language. An easy lesson to teach and master." Whenever the Experimenter observed this classroom, almost all students were on-task and the teacher was in full control.

School E_1, Middle Group. Instruction on SAPA lessons was generally very thorough. There was a good variety of large-group, small-group, and individual activities (e.g., large-group presentation for Lesson M, small-group experimentation during Lesson O, individual work at points during Lessons O and P). Typically, the introduction and vocabulary for each lesson were handled in 30 minutes or less. Students were not rushed during experimentation and small-group work. Following administration of a diagnostic test there was extensive review; each question was discussed and additional activities pursued for content areas that were not learned. The teacher utilized board-work and worksheets, when appropriate, to provide practice on important skills. For Lesson L (Guinea Pigs in a Maze) the teacher reported that the experimental animals were uncooperative; she modified the lesson by having children run through a classroom-constructed maze instead. For Lesson M, the final one in the sequence

for the middle group, the teacher noted: "I felt quite rushed getting through this lesson. I didn't think [the children] had really mastered it but gave the test anyway."

School E_1, Low Group. The teacher of this group experienced some difficulty in communicating effectively with the class. On occasions when the Experimenter was present, as many as one-half of the students seemed to be off-task. Almost all instruction was large-group, with the teacher lecturing and occasionally asking questions. There was little small-group and no individual work in this classroom. An inordinate amount of time was spent on the first unit of instruction (360 minutes) and the teacher was urged to move ahead to other lessons even if a substantial number of children had not achieved mastery on the diagnostic test. A major problem reported by the teacher was motivating the students to pay attention. After each diagnostic test, the teacher reviewed the test thoroughly, explaining the answer to each question in detail. As much as 90 minutes were spent for some lessons in reteaching. The teacher reported spending a substantial amount of time in teaching and reviewing key vocabulary words in the lessons.

School C_1, Heterogeneous Groups. In this school, three teachers taught roughly equal-size groups of children. There were children representing all three levels of pretest performance (high, middle, low) in each group. A round-robin technique was used; in the first round, one teacher took responsibility for Lesson L, one for Lesson M, and one for Lesson N. Lessons O, P, and Q were distributed in similar fashion for the second round of instruction. The round-robin technique had been used during the previous three semesters by the teachers in School C_1; thus, it was not a new instructional technique. One new feature of instruction, however, was the administration of SAPA-content tests (diagnostic tests) at points throughout the semester. Previously, as part of the longitudinal intervention study, students in the control schools received SAPA tests only at the beginning of a semester (pretest) and at the end of the semester (posttest). In the present study, the SAPA tests were also administered at the end of each lesson during the semester (in experimental and control schools alike) to control for testing effects. This had the undesirable consequence of introducing a new element into the pattern of instruction at Schools C_1 and C_2. The teachers at School C_1 felt that giving the SAPA tests immediately following instruction had the effect of motivating them and their students to cover the lessons more thoroughly. One teacher reported that his students were anxious to take the diagnostic tests to see how well they could do. Although no plan for formal feedback and review of test items was

established, teachers in School C_1 reported that students frequently initiated informal discussions of test items following the tests. Instruction on the SAPA lessons was generally thorough. There was primarily large-group discussion, with occasional small-group work involving experimental equipment. No tutoring or individual work was reported or observed in School C_1. One teacher reported using teacher-made worksheets to reinforce important concepts during instruction. Two teachers noted: "We really liked the idea of giving the tests after each lesson. The kids really saw some reason for paying attention during the lesson, because they knew the test was coming."

School E_2, High Group. In this group SAPA instruction was well planned and executed. Individual activity was encouraged and originality was reinforced. For each lesson, after 20-30 minutes of teacher-led introduction and vocabulary work, the students spent most of their time in small groups or worked individually. A significant degree of working in "pairs" was reported and observed. For example, during Lesson L the students created their own mazes and the teacher duplicated these for the class; children then tried "running" the mazes themselves and recorded how their maze time changed with practice. The teacher recorded in her log that "children write down the time it takes them and others challenge their time. Our baby gerbils made it difficult for us to run enough trials so I used information from the teacher guide for the children to graph." Throughout the SAPA lessons the teacher provided a number of self-made supplementary activities; these activities were not technically "enrichment"; however, the significant number of such "reinforcement" exercises increased the time spent by this high group to the point that high and middle students in School E_1 spent approximately the same time on instruction. The high group teacher was extremely conscientious and felt reluctant to move quickly from one lesson to the next without providing the supplementary exercises. Following initial instruction and diagnostic testing, the students reviewed the test questions and received additional instruction ranging from 10 to 30 minutes. An entire period was spent reteaching the content for Lesson N, while only 10 minutes was needed for most of the other lessons. The teacher noted: "Usually most [children] understood mistakes without too much discussion. Partners help for those having trouble."

School E_2, Middle Group. The teacher of this group was well-organized and efficient; she was always in control of the class and used a variety of techniques to interest and motivate the students. Large-group discussion, small-group work, and individual work were included in instruction. At the beginning of each SAPA lesson, the

teacher spent from 15 to 40 minutes teaching vocabulary and introducing the lesson content. Discussions were generally "discovery" oriented rather than lecture. During initial instruction children usually worked in small groups. The teacher provided a number of self-made worksheets to involve the students and to provide practice on such topics as interpolation, extrapolation, graphing, and report-writing. During observations by the Experimenter, nearly all students appeared to be on-task. Following initial instruction and diagnostic testing, the students discussed test items, and reteaching was accomplished in large-group discussion or individually (with teacher-made worksheets). In one instance, following the diagnostic test for Lesson P, the class was divided into two groups: all children who mastered the content were put on one side of the room and received a "rocket worksheet" stressing coordinates; the teacher worked with the non-mastery group of ten students, plotting points on the board until all children seemed to understand the coordinate system. The teacher of this group was very enthusiastic, and evidenced an ability to motivate students and to extend SAPA content to "real life" examples.

School E_2, Low Group. The teacher of this group was skilled and conscientious, but very troubled by the special demands of teaching a low section of students. Numerous entries in the teacher log indicated a painful awareness of the short attention span and limited motivation of these children . . . Lesson L: "The students showed great interest in the gerbils as personalities, but couldn't care less about SAPA. By the end of 15 minutes over half the kids had lost complete interest in the trial, the maze, or the animals. By the end of 24 minutes only about 25% were at all interested in the animals." Lesson M: "These kids need motivation as much or more as special instruction. They simply do not apply the ability they possess unless this application is demanded--even then some are in another world." Again for Lesson M, the following comments were entered in the log: "Whoever said this is a 'hands on' approach to science chose not to observe this lesson. If you think more time will solve the problem with slower students, you are mistaken. They don't have the attention span or interest necessary to drill on such abstract concepts. More time simply provides more time for inattention and boredom. In short, mastery of this lesson isn't relevant." Although the teacher of this group spent more total time on instruction than any other teacher in the entire study (see Table 6), the obvious conclusion is that time alone does not ensure mastery. Students in this group were organized for instruction primarily in large-groups and small-groups; no individual work was reported or observed. When the teacher was engaged in large-group instruction, a variety of aids were used: overhead projector, charts, boardwork, and laboratory

equipment. Following initial instruction (quite lengthy in some cases), the students reviewed the test questions and received additional instruction on what was not learned. This group spent eleven weeks on the six SAPA lessons.

School C_2, Heterogeneous Groups. In School C_2 instruction was departmentalized: a single teacher had responsibility for all science instruction at level six. Approximately six 30-minute periods were spent on each SAPA lesson. Initial instruction was always in a large group, with the teacher writing important points on the blackboard or asking questions. The predominant method of instruction was lecture. Only in cases where experimental equipment was used did the teacher divide the students into small groups of four to five. All activities specified in the SAPA teacher guides were taught; no supplementary activities were reported or observed. At the conclusion of instruction on a SAPA lesson, the diagnostic test was administered. As in all other experimental and control classrooms, the teacher read each question aloud and students were given ample time to mark their responses on the answer sheets. At the conclusion of testing all answer sheets and tests were collected, and the remaining time was spent in recreational reading. No test items were discussed and there was no reteaching of unmastered content. The teacher of this group was usually well organized and most students were observed to be on-task when the Experimenter visited.

PURPOSE AND HYPOTHESES

The present study was intended to examine the effects of a method of instruction based on instructional programming for the individual student. In this method, both differential grouping and differential pacing of students were applied to regular SAPA science instruction in the experimental treatment. Experimental Ss were placed into one of three groups (high, middle, or low) on the basis of cognitive entry characteristics (i.e., pretest level of achievement on the six SAPA lessons to be taught) and an attempt was made to secure mastery of the same science objectives by as many students as possible within each of the three groups. Teachers proceeded

through a sequence of SAPA lessons at pace determined by the needs of the group; high, middle, and low classes spent only as much time as needed by the students in the group before going ahead to the next lesson in the sequence. Specifically, the purpose of the present study was to determine the combined effects of the experimental treatment (i.e., grouping students for differentiated instruction on the basis of entry level of achievement, and varying the amount of instructional time) on three dependent variables (i.e., students' science achievement, number of students attaining mastery of lesson objectives, and attitudes of students toward science). The following research questions were posed:

1. Does instruction that is organized to promote mastery of objectives in accordance with instructional programming for the individual student facilitate children's learning of science, as measured by (a) achievement test performance, and (b) number of students who attain mastery?
2. Does instruction that is organized in accordance with instructional programming for the individual student have a significant effect upon student attitudes toward science?

In addition to the two basic research questions, there was a third question of interest to the writer.

3. Is there a relationship between the method of instruction (experimental treatment) and the instructional group to which a student belongs?

The possibility of an aptitude by treatment interaction, as expressed in question 3, influenced the choice of experimental design (i.e., 2 x 3 factorial) and analysis (i.e., two-way analysis of variance).

Hypotheses related to each of the research questions were as follows:

1a. Experimental groups receiving differentiated instruction designed to promote mastery of objectives will perform significantly better on achievement tests of SAPA-content knowledge than control groups receiving non-differentiated, large-group instruction.

1b. A significantly greater number of children will attain mastery of the science objectives in the experimental groups, taught in accordance with instructional programming for the individual student, than in the control groups.

2. Students in experimental schools, taught in accordance with instructional programming for the individual student, will score significantly higher than control students on an inventory of attitudes toward science, indicating a more positive affect toward science.

3. Taken as the null hypotheses, there will be no significant relationship (interaction) between the method of instruction and the instructional group to which a student belongs.

EXPERIMENTAL DESIGN AND DATA ANALYSES

A 2 x 3 (treatment x group) factorial design was used in the study. There were two levels (experimental, control) of the treatment factor and three levels (high, middle, low) of the group factor. Scores on the 60-item SAPA-content pretest were used to assign subjects to groups (raw scores 35 and above = high; raw scores 27-34 = middle; raw scores 26 and below = low). The design of the study is presented in Table 7.

Tests of Assumptions for ANCOVA

As originally proposed, data analyses for the present study included ANCOVA. The achievement scores on the 90-item SAPA-content posttest were to be analyzed by analysis of covariance, with the pretest scores serving as the covariate. During preliminary data analysis a significance test for equality of the twelve regression slopes was performed. The results of the test ($F = 1.87$; $df = 11, 211$; $p < .045$) indicated that the assumption of equality of regression slopes for analysis of covariance was violated. The individual slopes (posttest predicted from pretest) by group for each school are presented in Table 8. It can be seen that the regression slopes were quite disparate. In general, pretest score was the best predictor of posttest performance for students in the middle group at each school; pretest scores predicted posttest performance much less reliably for subjects in either of the other two groups, high and low. Exceptions to this general trend were the low group at School

TABLE 7

DESIGN OF THE STUDY

Group	Treatment	
	Experimental	Control
High		
Middle		
Low		

TABLE 8

REGRESSION SLOPES BY GROUP WITHIN EACH EXPERIMENTAL AND CONTROL SCHOOL

Group	School E_1	School C_1	School E_2	School C_2
High	0.46	1.11	0.48	1.60
Middle	1.00	1.45	1.03	-0.13
Low	-0.02	0.55	2.14	0.80

E_2 and the high and middle groups at School C_2.

Many sources (Glass, Peckham, & Sanders, 1972; Kirk, 1968; Hays, 1973; Scheffé, 1959; Winer, 1971) describe the very serious consequences of violating the assumption of homogeneity of regression slopes for analysis of covariance. According to Glass et al.,

> Interpretations of significant main effects in the presence of heterogeneous regression slopes in the analysis of covariance is similar to making main effect interpretations in the presence of interactions in the factorial analysis of variance set up. In the latter, however, the possibility of making meaningful statements about the independent variables is greater than in the former. This suggests that if heterogeneous regression slopes are suspected, the concomitant variable might be used as a blocking variable in the standard factorial analysis of variance. (1.72, pp. 275-276)

Due to the fact that analysis of covariance is not robust with respect to violation of the assumption of homogeneity of regression slopes, it was decided to abandon ANCOVA for use in the present study. A two-way analysis of variance was adopted instead. (See chapter four for tests of assumptions for ANOVA and results of data analysis.)

IV

RESULTS

PRETEST ACHIEVEMENT DATA

Inasmuch as analysis of covariance was inappropriate due to heterogeneity of regression slopes, it was decided to use an analysis of variance on the SAPA-content posttest scores. Before reporting the results of the ANOVA, however, a description of pretest scores for the various treatment groups will be given. Presented in Table 9 are numbers, means, and standard deviations for experimental and control groups on the 60-item SAPA-content pretest. As shown in the table, the pretest levels of performance for one pair of schools (E_1 and C_1) were unequal, favoring the experimental school. The pretest mean for all subjects in School E_1 was 31.40, roughly one-half standard deviation above the mean of 27.57 for all subjects in School C_1. Pretest means for the other pair of schools, E_2 and C_2, were nearly identical; the mean for School E_2 was 31.08 and the mean for School C_2 was 31.02. When data from all experimental subjects was combined, the mean was 31.23, somewhat less than one-half standard deviation above the mean of 29.26 for all control subjects combined.

It can be seen from Table 9 that all experimental groups (high,

TABLE 9

NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS FOR EXPERIMENTAL AND CONTROL GROUPS ON COMBINED SAPA-CONTENT TESTS L-Q: 60-ITEM PRETEST

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	39.70	37.08	41.33	37.17	40.59	37.13
S.D.	4.89	3.64	5.82	3.79	5.42	3.67
Middle						
Number	20	16	18	24	38	40
Mean	31.45	29.75	31.00	30.63	31.24	30.28
S.D.	3.25	2.84	3.91	2.00	3.54	2.37
Low						
Number	20	27	23	12	43	39
Mean	23.05	21.70	20.43	22.58	21.65	21.97
S.D.	4.59	4.29	4.49	5.04	4.67	4.48
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	31.40	27.57	31.08	31.02	31.23	29.26
S.D.	8.06	7.31	10.16	6.37	9.18	7.05

middle, and low) in School E_1 scored higher on the pretest than their respective groups in the matched control school. This was also true for experimental and control groups in the E_2,C_2 pair, except for low students in School E_2, who scored lower on the pretest than the low group in School C_2. When data from combined experimental schools (E_1E_2) was compared to combined control schools (C_1C_2), the greatest difference in pretest performance was observed between experimental and control subjects in the high group; subjects in the middle group were also different, but to a lesser degree; and pretest performance was nearly equal between experimental and control subjects in the low group. In the case of the high and middle groups, the differences favored the experimental schools; for the low group, however, the mean for control subjects was slightly higher than for experimental subjects. (Note: Pretest results are reported separately for each 10-item SAPA-content test in Appendix D).

POSTTEST ACHIEVEMENT DATA

One of the research questions posed in the present study was the following: Does instruction that is organized to promote mastery of objectives in accordance with the IPM facilitate children's learning of science, as measured by (a) achievement test performance, and (b) number of students who attain mastery? In this section, the analysis of data related to part a will be reported. (Note: Posttest results are reported separately for each 10-item SAPA-content test in Appendix E.)

Results of ANOVA

The numbers of subjects, means, and within-cell standard deviations for experimental and control groups on the 90-item SAPA-content posttest are presented in Table 10. Results of the two-way analyses of variance performed on the posttest data are given in three tables: Table 11 presents the results for School E_1 paired with School C_1; Table 12 presents the results for School E_2 paired with C_2; and Table 13 presents the results for the combined experimental schools E_1E_2 paired with the combined control schools C_1C_2.

As indicated in Table 11, for School E_1 paired with School C_1 there was no significant main effect for Treatment. The calculated F values on the main effects did indicate a significant difference due to Group ($F = 87.01$; $df = 2,110$; $p < .0001$) and a significant Treatment x Group interaction ($F = 6.45$; $df = 2,110$; $p < .002$). An inspection of Table 10 reveals that the reason for the nonsignificant Treatment effect was the high level of performance at School C_1 ($\bar{X} = 65.52$, S.D. = 11.28) relative to the paired experimental school E_1 ($\bar{X} = 69.05$, S.D. = 14.38). The difference between means was in the predicted direction, with experimental Ss attaining higher scores than control Ss. It can also be seen from Table 10 that the significant main effect for the Group factor was due to the differential performance on the posttest of the high, middle, and low groups. Inspection of Figure 5 reveals that the difference in posttest achievement was slightly greater between low and middle groups than

TABLE 10

NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS FOR EXPERIMENTAL AND CONTROL GROUPS ON COMBINED SAPA-CONTENT TESTS L-Q: 90-ITEM POSTTEST

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	80.20	75.92	81.92	65.47	81.14	69.85
S.D.	5.85	7.37	5.15	11.11	5.48	10.91
Middle						
Number	20	16	18	24	38	40
Mean	74.90	69.44	72.39	58.13	73.71	62.65
S.D.	6.54	7.47	7.54	9.91	7.05	10.53
Low						
Number	20	27	23	12	43	39
Mean	52.05	58.19	55.61	49.75	53.95	55.59
S.D.	9.68	9.62	15.93	11.27	13.36	10.75
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	69.05	65.52	69.97	58.71	69.53	62.18
S.D.	14.38	11.28	15.53	11.93	14.93	12.05

TABLE 11

RESULTS OF TWO-WAY ANALYSIS OF VARIANCE ON 90-ITEM SAPA-CONTENT POSTTEST FOR SCHOOL E_1 PAIRED WITH SCHOOL C_1

Source of Variation	df	MS	F	p less than
Treatment	1	39.79	0.61	.435
Group	2	5646.58	87.01	.0001
Treatment x Group	2	418.46	6.45	.002
Error	110	64.90		

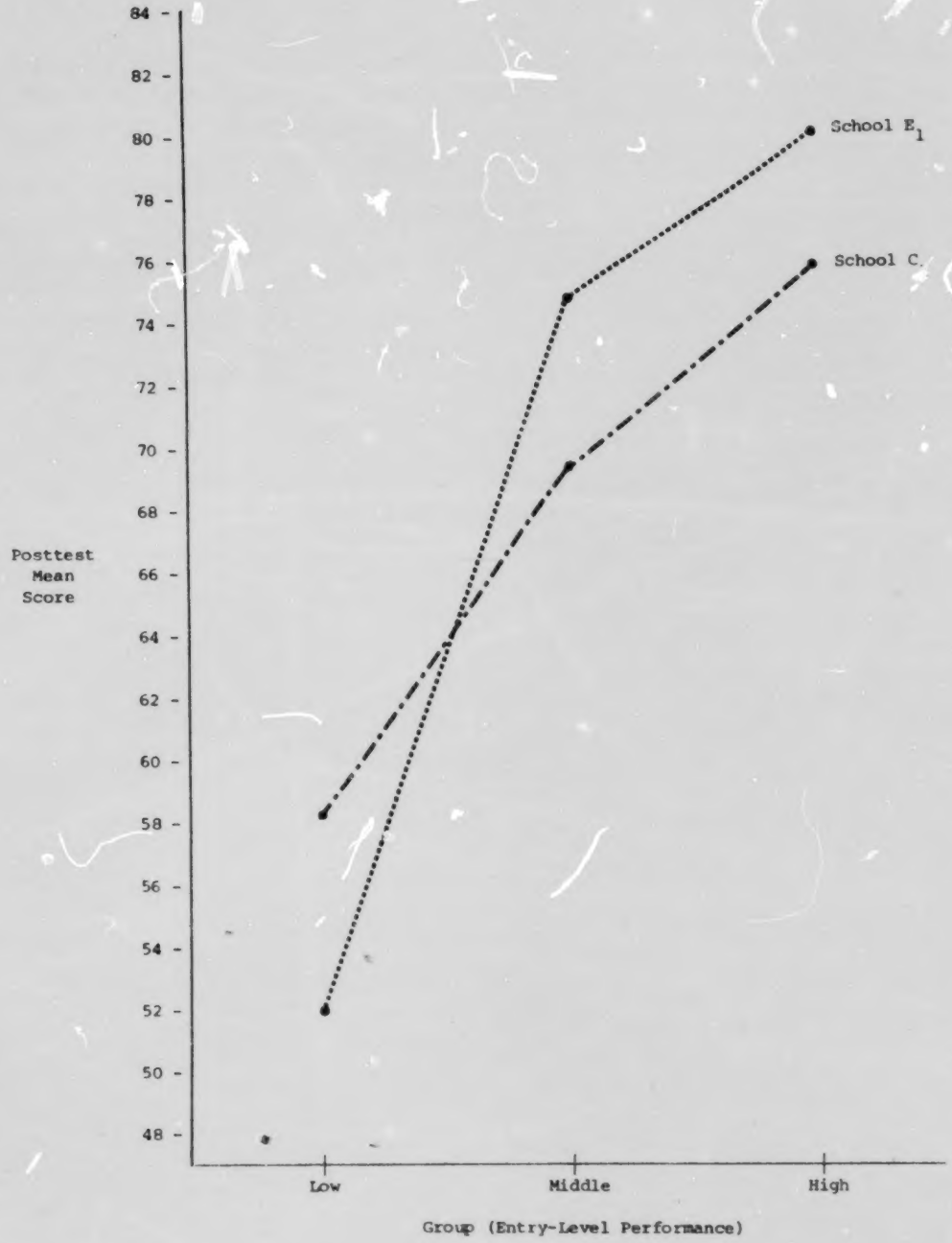


Figure 5. Mean scores on 90-item SAPA-content posttest by group for School E_1 and School C_1.

between middle and high groups. Figure 5 also illustrates the source of the significant Treatment x Group interaction: Ss in the high and middle groups at the experimental school scored higher than Ss in the high and middle groups at the control school; however, students in the low group at the control school actually scored higher than low Ss at the experimental school. To summarize the results for the E_1,C_1 pair of schools, there was no significant Treatment effect, while there was a significant effect for Group (high > middle > low) and a significant interaction effect, that of Treatment x Group. The interaction effect was attributable to the performance of Ss in the low groups, where low children in the experimental condition failed to perform as well as children in the control condition. From Table 10 it can be seen that the mean (58.19) for the low group at C_1 was higher than for any other low group, including the other experimental school, E_2.

As indicated in Table 12, for School E_2 paired with School C_2 there were significant main effects for Treatment ($F = 36.44$, $df = 1,113$, $p < .0001$) and for Group ($F = 34.55$, $df = 2,113$, $p < .0001$). From Table 10 it can be observed that all means were in the predicted direction: experimental Ss in each group scored higher than control Ss. Figure 6 illustrates that the largest difference between means occurred for the high groups; differences between means decreased systematically from middle to low groups. The significant main effect for the Group factor was due to the differential perfor-

TABLE 12

RESULTS OF TWO-WAY ANALYSIS OF VARIANCE ON 90-ITEM SAPA-CONTENT POSTTEST FOR SCHOOL E_2 PAIRED WITH SCHOOL C_2

Source of Variation	df	MS	F	p less than
Treatment	1	4163.36	36.44	.0001
Group	2	3947.79	34.55	.0001
Treatment x Group	2	269.40	2.36	.099
Error	113	114.25		

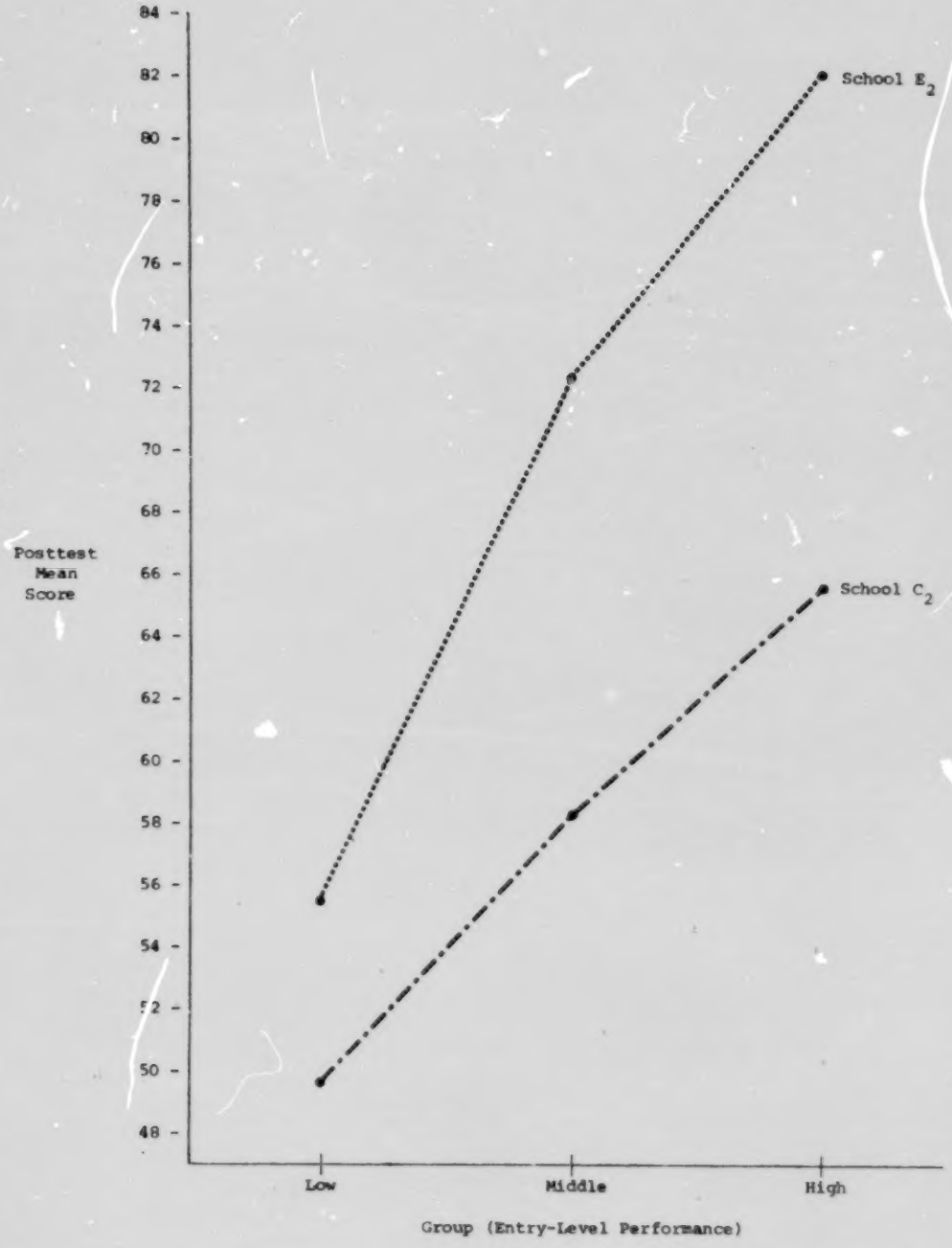


Figure 6. Mean scores on 90-item SAPA-content posttest by group for School E_2 and School C_2.

mance on the posttest of the high, middle, and low groups; mean scores were in the direction high > middle > low. For School E_2 paired with School C_2, there was no significant interaction between Treatment and Group. To summarize the results for the E_2C_2 pair of schools, there were significant main effects for both Treatment and Group, and no significant Treatment x Group interaction.

Presented in Table 13 are the results of the two-way analysis of variance for the combined experimental schools E_1E_2 paired with the combined control schools C_1C_2. The calculated F values on the main effects indicated significant differences due to Treatment (F = 27.61, df = 1,229, $p < .0001$) and to Group (F = 85.06, df = 2,229, $p < .0001$); there was also a significant Treatment x Group interaction (F = 10.86, df = 2,229, $p < .0001$). An inspection of Table 10 reveals that means were in the expected direction (experimental > control) only for the high and middle groups; the mean performance of the combined experimental Ss in the low group at E_1 and E_2 was lower than the level of achievement associated with low Ss in the combined control schools. As shown in Table 10, the mean for the combined control schools in the low category was heavily weighted by the Ss in School C_1; there were over twice as many subjects in the low group at School C_1 (n = 27) than at School C_2 (n = 12). For this reason the combined C_1C_2 mean of 55.59 was more typical of the C_1 mean (58.19) than the C_2 mean (49.75). The unexpected performance of experimental and control children in the low groups (con-

TABLE 13

RESULTS OF TWO-WAY ANALYSIS OF VARIANCE ON 90-ITEM SAPA-CONTENT POSTTEST FOR COMBINED EXPERIMENTAL SCHOOLS E_1E_2 WITH COMBINED CONTROL SCHOOLS C_1C_2

Source of Variation	df	MS	F	p less than
Treatment	1	2763.23	27.61	.0001
Group	2	8613.56	86.06	.0001
Treatment x Group	2	1086.95	10.86	.0001
Error	229	100.09		

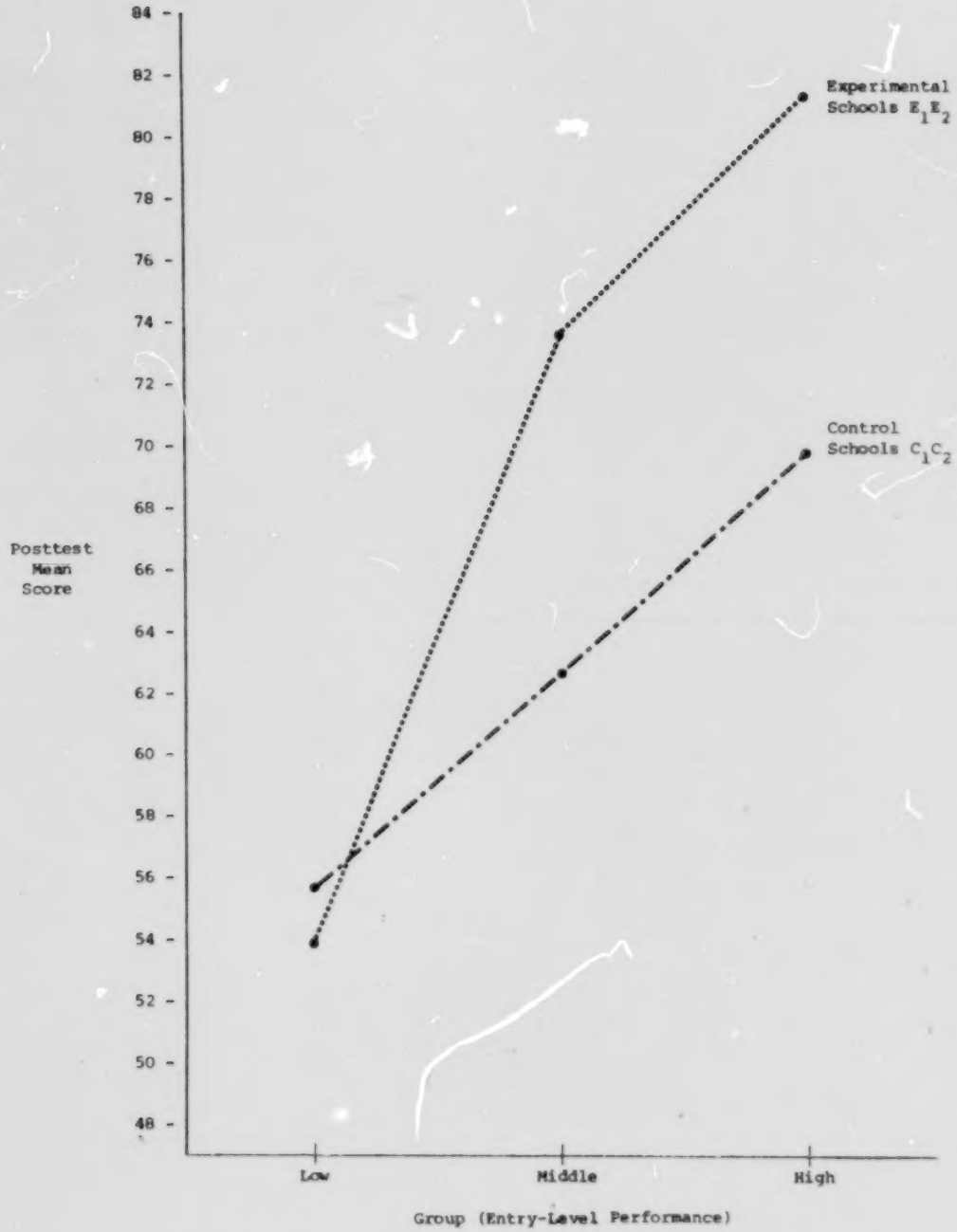


Figure 7. Mean scores on 90-item SAPA-content posttest by group for combined experimental schools E_1E_2 and combined control schools C_1C_2.

trol > experimental) can be isolated as the reason for the significant Treatment x Group interaction. As illustrated in Figure 7, there was a differential effect for experimental treatment across groups, such that for some groups (high and middle) the treatment had a facilitative effect, while for another group (low) the treatment did not facilitate performance on the posttest. The significant Group effect was due to the differential performance of the high, middle, and low groups on the dependent measure. To summarize the results for the combined experimental schools paired with the combined control schools, there were significant main effects for both Treatment and Group, as well as a significant Treatment x Group interaction.

Tests of Assumptions for ANOVA

It can be observed from Table 10 that the variances associated with reported means are quite different from cell to cell; for example, the variance for the high group in School E_2 is 26.52 compared to a variance of 253.76 for the low group in the same school. Because of the suspected heterogeneity of variances, the Bartlett Test for homogeneity of population-error variances (Kirk, 1968, p. 61) was performed. For the six cells in the E_1,C_1 comparison (three groups at each of two schools) the test was nonsignificant (B = 8.09, df = 5, p < .25), indicating homogeneity of variances. However, for the two other comparisons, E_2 with C_2 and combined E_1E_2 with combined C_1C_2, the null hypothesis of equal variances was rejected. For the

six cells in the E_2,C_2 comparison, $B = 28.30$, $df = 5$, $p < .001$; in the case of combined experimental schools paired with combined control schools, $B = 38.25$, $df = 5$, $p < .001$. Thus, for two of the three analyses of variance, the assumption of homogeneity of variances was violated.

As indicated by Glass et al., the major effects of violations of assumptions for analysis of variance are now well known, after nearly four decades of research. "The relevant question is not whether ANOVA assumptions are met exactly, but rather whether the plausible violations of the assumptions have serious consequences on the validity of probability statements based on the standard assumptions" (Glass, Peckham, & Sanders, 1972, p. 237). It is widely held that the F distribution is robust with respect to violation of the assumption of homogeneity of variances provided that the number of observations in the samples is equal (Glass et al., 1972; Hays, 1973; Kirk, 1968; Scheffé, 1959; Winer, 1962). However, for samples of unequal size, violation of the homogeneity assumption can have a serious effect on tests of significance. Reprinted in Glass et al. is a table constructed by Box (1954) showing the effect of heterogeneous variances on the probability of Type-I error in an analysis of variance. Box shows how a nominal significance level of .05 can can be affected for various samples of unequal n's and unequal variances. When n's are unequal and variances heterogeneous, the actual significance level may greatly exceed or be greatly exceeded by

the nominal significance level, depending upon the relationship between the n's and variances. Specifically, when samples with smaller n's come from populations with larger variances, the actual significance level may exceed the nominal level (e.g., the probability of a Type-I error may actually be .17 when the experimenter thinks it is .05 in a case where $n_1 = 9$, $n_2 = 5$, $n_3 = 1$, and the variances are in the ratio 1:1:3). Conversely, when n's and variances are directly proportional (e.g., $n_1 = 3$, $n_2 = 5$, $n_3 = 7$, and the variances are in the ratio 1:2:3) the actual probability of a Type-I error may be less than the nominal level (e.g., $p < .04$ when the experimenter thinks it is .05). In the present study the heterogeneous variances are not directly proportional; neither, however, are they systematically related in such a way that smaller n's are always associated with larger variances. The entry in Box's (1954) table that is most similar to the data in the present study is the following: $n_1 = 3$, $n_2 = 9$, $n_3 = 3$, with variances in the ratio 1:2:3. In this instance, the actual probability of a Type-I error is .056 when the nominal level of significance is .05. It can be seen, therefore, that in some cases when sample sizes are not grossly different (even though they are unequal) and variances are heterogeneous, the difference between actual and nominal levels of significance is not great.

In the present study, whenever F-ratios have indicated significant effects, the probabilities for making Type-I errors have been very small ($p < .0001$ or in one case .002). It seems reasonable to

conclude that the main effects and interactions reported in this chapter are probably significant despite the violation of the homogeneity of variances assumption and the unequal cell sizes. In any event, the reader would be wise to interpret with caution the results reported herein; any violation of a basic assumption underlying the analysis of variance test ought to be taken into account when viewing the results of the analysis.

POSTTEST MASTERY DATA

As stated previously, one of the research questions posed in the present study was the following: Does instruction that is organized to promote mastery of objectives in accordance with the IPM facilitate children's learning of science, as measured by (a) achievement test performance, and (b) number of students who attain mastery? In this section, the analysis of data related to part b will be reported.

Presented in Table 14 are a number of 2 x 2 contingency tables, indicating the number of students who attained mastery (at the 80% level) and the number of students who failed to attain mastery on the SAPA-content posttest. The six 15-item tests associated with the six SAPA units of instruction were combined to form the 90-item posttest; the criterion of mastery was set at 72 or more items answered correctly. A series of 12 Chi-square tests of association were performed on the data, and the results are included in Table 14; in all cases, Yates' correction for continuity (Hays, 1973, p. 735)

TABLE 14

NUMBERS OF STUDENTS AT MASTERY AND NON-MASTERY AND RESULTS OF CHI-SQUARE TESTS BY GROUP: SAPA-CONTENT POSTTESTS L-Q COMBINED

		E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High Group	Mastery	19	8	23	7	42	15
	Non-Mastery	1	5	1	11	2	16
		$\chi^2 = 3.89^*$		$\chi^2 = 13.67^{***}$		$\chi^2 = 19.58^{***}$	
Middle Group	Mastery	16	7	11	0	27	7
	Non-Mastery	4	9	7	24	11	33
		$\chi^2 = 3.61$		$\chi^2 = 16.84^{***}$		$\chi^2 = 20.60^{***}$	
Low Group	Mastery	1	2	4	0	5	2
	Non-Mastery	19	25	19	24	38	49
		$\chi^2 = 0.07$		$\chi^2 = 2.60$		$\chi^2 = 1.05$	
Combined Groups Within School	Mastery	36	17	38	7	74	24
	Non-Mastery	24	39	27	47	51	86
		$\chi^2 = 9.10^{**}$		$\chi^2 = 24.07^{***}$		$\chi^2 = 32.11^{***}$	

$^*p < .05$
$^{**}p < .01$
$^{***}p < .001$

was used since there was only 1 degree of freedom. An inspection of the table reveals that in the high group for all three experimental/control comparisons (E_1 with C_1, E_2 with C_2, and E_1E_2 with C_1C_2) there were significantly more Ss at mastery in the experimental condition than in the control condition ($p < .05$ for E_1 with C_1; $p < .001$ for E_2 with C_2 and for E_1E_2 with C_1C_2). In the middle group there were significantly more experimental Ss at mastery in the E_2,C_2 pair of schools ($p < .001$) and in the combined experimental schools paried with the combined control schools ($p < .001$); there was no difference, however, between the number of children at mastery and non-mastery in School E_1 paired with School C_1. In the low group differences between the number of Ss at mastery and non-mastery were not significant for any of the three experimental/control comparisons. Figures 5, 6, and 7 graphically illustrate the similarity in performance of low-group children. The mastery data confirm the fact that in this study there was a significant interaction between experimental treatment and group (based on entry level of achievement): the treatment generally facilitated learning for high and middle Ss while proving ineffective for Ss in the low group. When mastery/non-mastery performance was compared using scores from all Ss within a school (i.e., combined across groups), there was a significant difference between the number of children at mastery and non-mastery for all experimental/control contrasts: E_1 with C_1 ($p < .01$); E_2 with C_2 ($p < .001$); and E_1E_2 with C_1C_2 ($p < .001$). (Note:

Numbers of students at mastery and non-mastery and results of Chi-square tests by group are reported separately for each 15-item SAPA-content test in Appendix F.)

Table 15 presents the number and percentage of students in each school (combined across groups) attaining mastery (80% level) on each of the six units of SAPA instruction (L, M, N, O, P, and Q). An inspection of the table reveals that the percentage of students at mastery varied widely from lesson to lesson. For example, in the combined control schools, only 19.1% of the Ss attained mastery on SAPA Lesson N, while 50% attained mastery on Lesson P. This suggests that lesson content was a variable of some significance in the present study. From the table it can be seen that the mean number of students attaining mastery across lessons was always greater for the experimental schools than for the control. This difference (in the expected direction) was less pronounced for the E_1,C_1 pair of schools than for the E_2,C_2 pair. In the combined experimental schools the mean number of Ss attaining mastery of SAPA lessons was equivalent to 61.86% of all students in Schools E_1 and E_2; in the combined control schools the mean number of Ss attaining mastery was equivalent to 38.64% of the C_1 and C_2 population.

POSTTEST AFFECTIVE DATA

The Science Attitude Inventory (see Appendix C) was used to assess the affective consequences of the experimental treatment. It was hypothesized that children in the experimental schools receiving

TABLE 15

NUMBER AND PERCENTAGE OF STUDENTS IN EACH SCHOOL ATTAINING MASTERY (80% LEVEL) ON INDIVIDUAL SAPA UNITS OF INSTRUCTION

School Comparison		Students Attaining Mastery per Unit of Instruction						Mean Number of Students Attaining Mastery Across Units of Instruction	Mean Number of Students Attaining Mastery as a % of School Population
		L	M	N	O	P	Q		
E_1 ($\underline{n}$ = 60)	N:	42	34	33	37	42	27	35.83	59.72%
	%:	70.0	56.7	55.0	61.7	70.0	45.0		
C_1 ($\underline{n}$ = 56)	N:	31	19	15	34	33	21	25.50	45.54%
	%:	55.4	33.9	26.8	60.7	58.9	37.5		
E_2 ($\underline{n}$ = 65)	N:	50	39	36	46	41	37	41.50	63.85%
	%:	76.9	60.0	55.4	70.8	63.1	56.9		
C_2 ($\underline{n}$ = 54)	N:	23	14	6	20	22	17	17.00	31.48%
	%:	42.6	25.9	11.1	37.0	40.7	31.5		
E_1E_2 ($\underline{n}$ = 125)	N:	92	73	69	83	83	64	77.33	61.86%
	%:	73.6	58.4	55.2	66.4	66.4	51.2		
C_1C_2 ($\underline{n}$ = 110)	N:	54	33	21	54	55	38	42.50	38.64%
	%:	49.1	30.0	19.1	49.1	50.0	34.5		

instruction in accordance with instructional programming for the individual student would exhibit a more positive attitude toward science than children in control schools receiving non-differentiated instruction. As part of the longitudinal intervention study, the 24-item SAI was administered in all four schools during December 1976; the results of that administration were used to provide a baseline estimate of attitude in the present study. Twelve items on the 24-item inventory were scored, according to the procedure outlined in chapter three; the highest possible score was 48 (extremely positive attitude) and the lowest possible score was 12 (extremely negative attitude).

The baseline mean score on the SAI for experimental Ss ($\bar{X}$ = 36.86, S.D. = 8.73) was not significantly different (t = .68, df = 220, $p < .50$) from the mean score for control Ss ($\bar{X}$ = 36.08, S.D. = 8.50). Since baseline attitude scores were not different between experimental and control students, it was decided to use ANOVA for the analysis of posttest attitude scores. The SAI posttest was administered to students upon completion of instruction on the six SAPA lessons and prior to administration of the 90-item SAPA-content posttest. Table 16 presents the means and standard deviations for experimental and control school combinations on the affective inventory. Bartlett's Test for homogeneity of variances revealed that the variances were significantly different for two of the three comparisons: School E_1 paired with School C_1 (B = 4.88,

TABLE 16

MEANS AND STANDARD DEVIATIONS FOR EXPERIMENTAL AND CONTROL SCHOOLS ON SCIENCE ATTITUDE INVENTORY POSTTEST

	School Comparison					
	E_1 (N=57)	C_1 (N=50)	E_2 (N=61)	C_2 (N=53)	E_1E_2 (N=118)	C_1C_2 (N=103)
Mean	36.70	34.00	35.54	36.36	36.10	35.21
S.D.	7.81	10.62	8.29	9.11	8.05	9.89

df = 1, p < .05) and combined experimental E_1E_2 paired with combined control C_1C_2 (B = 4.61, df = 1, p < .05). For the E_1 with C_1 comparison the variances were not significantly different (B = .49, df = 1, p < .50). The reader is referred to pages 108 - 110 for a discussion of the consequences of heterogeneous variances when ANOVA is employed.

Tables 17, 18, and 19 present the results of the one-way analyses of variance on SAI data for School E_1 paired with School C_1, School E_2 paired with School C_2, and combined schools E_1E_2 paired with combined schools C_1C_2, respectively. No significant main effects due to Treatment were found for any of the three comparisons. Thus, it can be concluded that there were no differences in attitude toward science between experimental and control students at the end of the study.

TABLE 17

RESULTS OF ONE-WAY ANALYSIS OF VARIANCE ON SCIENCE ATTITUDE INVENTORY POSTTEST FOR SCHOOL E_1 PAIRED WITH SCHOOL C_1

Source of Variation	df	MS	F	p less than
Treatment	1	194.43	2.28	.134
Within Groups	105	85.14		

TABLE 18

RESULTS OF ONE-WAY ANALYSIS OF VARIANCE ON SCIENCE ATTITUDE INVENTORY POSTTEST FOR SCHOOL E_2 PAIRED WITH SCHOOL C_2

Source of Variation	df	MS	F	p less than
Treatment	1	18.95	0.25	.617
Within Groups	112	75.33		

TABLE 19

RESULTS OF ONE-WAY ANALYSIS OF VARIANCE ON SCIENCE ATTITUDE INVENTORY POSTTEST FOR COMBINED EXPERIMENTAL SCHOOLS E_1E_2 PAIRED WITH COMBINED CONTROL SCHOOLS C_1C_2

Source of Variation	df	MS	F	p less than
Treatment	1	43.38	0.54	.463
Within Groups	219	80.13		

V

DISCUSSION AND CONCLUSIONS

The purpose of the present study was to examine the effects of adapting instruction in elementary science to students' entering achievement levels, based on instructional programming for the individual student (Klausmeier, 1977). In this method, both differential grouping and differential pacing of students were practiced, and an attempt was made to secure mastery of the same science objectives by as many students as possible within each of three relatively homogeneous groups.

EXPERIMENTAL EFFECTS: SCIENCE ACHIEVEMENT

It was hypothesized that Ss in the experimental condition (receiving differentiated instruction designed to promote mastery of objectives) would score significantly higher on tests of SAPA-content knowledge than control Ss (receiving non-differentiated instruction in relatively heterogeneous groups). The predicted treatment effect was observed in two of the three comparisons performed: (1) Ss in the experimental condition (schools E_1 and E_2 combined) scored significantly higher ($p < .0001$) on the SAPA-content posttest than Ss in the control condition (schools C_1 and C_2 combined); and (2) Ss in experimental school E_2 scored significantly higher ($p < .0001$) than

Ss in the matched control school, C_2. In the case of School E_1 paired with School C_1, there was no significant difference in posttest scores although the means were in the predicted direction (experimental > control). Ignoring for the moment the effects of the Group factor and Treatment by Group interaction effects, it can be concluded that in general the experimental treatment facilitated students' performance on the measure of science achievement.

An unexpected result in the present study was the exceptional performance of students at one of the control schools (C_1). Posttest means for the two experimental schools were less than one point apart; however, the posttest mean for School C_1 (65.53) was much higher than the mean for the other control school, C_2 (58.71). The surprisingly high scores of children at control school 1 may have been due to several factors. First, certain instructional conditions at School C_1 may have resulted in the facilitation of learning. In particular, there was some review of lesson content prior to diagnostic testing and the test items were sometimes discussed in class. This unintended use of the SAPA-content tests in the control condition may have provided teachers in School C_1 with certain information about individual differences that formally or informally could have been used to adapt instruction. As indicated in the descriptive summary on page 85, the teachers at control school 1 felt that administration of the SAPA-content tests at intervals throughout the semester had resulted in heightened student motivation. The note-

worthy departure from intended instructional conditions at School C_1 was coupled with the use of an efficient method of instruction and division of teaching responsibilities. For several semesters, the teachers of level six science at School C_1 had employed the round-robin technique by which division of responsibilities for the various lessons was accomplished. According to the teachers, this technique permitted more thorough preparation for instruction and served to motivate the students by exposing them to different styles of teaching. Regardless of what claims may be valid for the technique itself, the writer has observed that the staff was quite pleased with the round-robin arrangement; such acceptance and enthusiasm ought certainly to be taken into account when assessing the impact of the school's total instructional program.

It is the writer's judgment that of the explanations offered above, the factor most likely to have contributed to the high performance of C_1 subjects is the motivational effect of diagnostic testing throughout the semester. This testing was clearly a new feature of instruction in that it was introduced after five consecutive semesters of pre- and posttesting only at the beginning and at the end of each semester. While diagnostic testing was a new feature of the instructional environment in School C_2 as well, the C_2 teacher felt strongly that his students were being "overtested" and was not likely to have utilized the tests as a motivational factor. At School C_1, where the performance of the students was unexpectedly

high, the teachers reported: "We really liked the idea of giving the tests after each lesson. The kids really saw some reason for paying attention during the lesson, because they knew the test was coming." It is thought by the writer that the introduction of a new testing arrangement in the control schools may have resulted in a Hawthorne Effect at School C_1: the introduction of a novel set of research procedures often leads to significant changes in behavior due solely to participation in the research.[1] A weakness of the present study was failure to pre- and postassess the motivational level of all students, since motivation is such an important variable in the classroom setting.

Aside from the determination of main effects in the analysis of data, one of the research questions to be answered in the present

---

[1]It is interesting to note that all four schools in the present study have exhibited a shift in achievement performance due to participation in the longitudinal intervention study during the past three years. At the onset of experimentation in 1974, schools E1, E2, C1, and C2 were ranked in the lower third of all schools in the district, as judged by the administration of the school system. At the beginning of the 1976-77 school year, the central office science coordinator administered a 50-item test of SAPA-content knowledge to children in all fourteen elementary schools in the district. The pretest was intended to measure baseline performance from which to compare end-of-year data. In the absence of any effect due to past association with the experimental procedures, one would expect students in schools E1, E2, C1, and C2 to score no higher on the pretest than students in other schools; in fact, one might predict that the means would fall into the lower third based on data available in 1974. Of the fourteen schools tested, however, schools E1, E2, C1, and C2 ranked first, second, fourth, and third, respectively. Thus, even "control" schools in the present study have been influenced by participation in the research effort.

study was the following: Is there a relationship between the method of instruction (treatment condition) and the group to which a student belongs? This question influenced the choice of experimental design (i.e., 2 x 3 factorial) and analysis (i.e., two-way analysis of variance). As reported in chapter four, there was a significant group effect ($p < .0001$) for each of the three experimental/control comparisons: E_1 paired with C_1, E_2 paired with C_2, and combined experimental paired with combined control. A significant treatment by group interaction was also obtained for two comparisons: E_1 paired with C_1, and combined experimental paired with combined control.

The interaction between treatment and group indicated that the experimental treatment was differentially effective across groups: experimental procedures facilitated learning for the high and middle students, but not for the low students. The interaction effects (see Table 10, Figure 5, and Figure 7) occurred mainly because Ss in the low group at School C_1 attained unexpectedly high posttest scores. In fact, the posttest mean for C_1 (58.19) exceeded that of any other low group in the study (52.05, 55.61, and 49.75 for schools E_1, E_2, and C_2, respectively). Based on classroom observations of low-group instruction at School E_1 (see page 85), there may also have been some slight depression of E_1 scores due to teacher or classroom effects: in the course of monitoring instruction the writer noted frequent inattention and off-task behavior in the E_1 low group. The exact contribution of each factor noted above (i.e., facilitation of

performance at School C_1; possible teacher or classroom effects at School E_1) cannot be determined.

The differential achievement of homogeneous groups of students is not an uncommon finding. In the large-scale Utah Study (Borg, 1956) research was directed to the question of achievement differences in ability grouped and random grouped classrooms. One school district, District A, employed ability grouping; another school district, District R, employed random or heterogeneous grouping. In general, the goal of District A was to adapt curricular materials to the different ability levels and concurrently adjust the rate of presentation to the level of the child. In comparing the achievement of District A and District R pupils within each of three ability classifications (superior, average, and slow) Borg (1965) reached the following conclusions:

1. For superior students, the evidence appeared to favor ability grouping. Of 47 statistical comparisons between superior pupils in the two districts, 15 comparisons favored District A, three favored District R, and the remainder were not significant.
2. For average students, the evidence slightly favored the ability-grouped sample. Of the 47 statistical comparisons between average pupils, 13 comparisons favored District A, six favored District R, and the remainder were not significant.

3. For slow students in the two districts, there were four significant differences favoring District A pupils, and seven favoring District R pupils, with the remaining 36 differences being non-significant. These data, according to Borg, seemed to suggest that in terms of achievement the two grouping methods were about equally effective for slow pupils. "If any advantage is present, it would appear from the data of this research that the random grouping system is slightly more effective for the slow pupil" (1965, p. 34).

Other evidence (Goldberg, Passow, & Justman, 1966) also suggests the possible advantage of homogeneous grouping for high (and sometimes middle) ability students but not for low ability pupils.

Results of the present study are in line with the literature suggesting no apparent advantage in terms of achievement for the homogeneous grouping of low ability students. It may even be that the unexpectedly high achievement of control students in School C_1 was related to heterogeneous grouping for instruction. There is no way in which this possibility can be examined post hoc; due to the combination of homogeneous grouping and differentiated instruction in the experimental treatment, no valid statement of the functional realtionship between one or the other (i.e., grouping or differentiated instruction) and achievement can be made.

In the present study the interpretation of group effects is also

confounded because "group" was operationally different for experimental and control schools. Whereas entry level of achievement (pretest scores) determined the assignment of Ss to instructional groups (high, middle, low) in the experimental schools, the entry level of Ss had no instructional implication for control Ss who were taught in heterogeneous classes. At schools C_1 and C_2 the pretest scores were used only for the purpose of data analysis. This differential application of the pretest scores--grouping for instruction in the experimental condition, while grouping only for data analysis in the control condition--may have been a conceptual error.

EXPERIMENTAL EFFECTS: MASTERY OF OBJECTIVES

It was hypothesized that a significantly greater number of students would attain mastery of lesson objectives in the experimental groups than in the control groups. Results of Chi-square tests on the mastery data in two of three comparisons (E_2 with C_2, and E_1E_2 with C_1C_2) clearly indicated that significantly ($p < .001$) more high and middle students mastered lesson objectives in the experimental condition than in the control condition (see Table 14). As was the case for achievement scores, the low ability students did not perform differently between experimental and control conditions. Once again, the results are in line with the research literature suggesting no clear advantage of grouping methods or differential pacing for low ability students.

When the group factor was discounted by collapsing mastery data

across groups, it was found that for all three experimental/control comparisons the number of experimental students attaining mastery (80% level) on the 90-item SAPA-content posttest was significantly greater than the number of control students (see Table 14.)

EXPERIMENTAL EFFECTS: ATTITUDES TOWARD SCIENCE

It was hypothesized that children in the experimental treatment would evidence more positive attitudes toward science at the conclusion of the study. On the Science Attitude Inventory posttest there were no significant differences between experimental and control subjects for any of the experimental/control comparisons. (All groups expressed a generally positive attitude.) One explanation for the lack of significant differences is the questionable validity of the assessment instrument; while the SAI was highly reliable (alpha coefficient for the composite = .91), it may not have been high in content validity. That is, the "true" attitudes of the students toward science may not have been accurately reflected in the SAI scores. It is the writer's opinion that many of the items in the attitude inventory tap something other than attitude toward science, e.g., attitude toward school in general, attitude toward classroom environment, attitude toward teacher, attitude toward self. Due to the questionable validity of the SAI instrument, no firm conclusions can be drawn relative to the affective consequences of the experimental treatment.

LIMITATIONS OF THE STUDY

Among the limitations of the study already mentioned have been the failure to assess motivational levels of the Ss and the problem of operationally different "groups" in the experimental and control conditions. According to the Instructional Programming Model (Klausmeier, 1976, 1977) children's motivational levels are assessed through observations by teachers during the school year (Klausmeier & Goodwin, 1975). However, in the present study no systematic effort was made to assess each child's level of motivation and adapt instruction accordingly. Another learner characteristic not assessed in the present study was learning style. In the IPM this characteristic is also taken into account when adapting instruction to the individual student. Whether a child learns more effectively through printed materials, audio-visual materials, or teacher instruction, whether the child prefers working alone, in pairs, or in small groups, and whether a child is largely self-directed or in need of external direction are all taken into consideration when adapting instruction to the individual. Because motivational level and learning style were not dealt with in the present study, the results may be limited in generalizability.

It is conceivable that the teacher variable may have been yet another threat to the internal validity of the present study. Even though teachers were randomly assigned to groups (high, middle, low) in the experimental schools, it is likely that some systematic bias

resulted from a teacher consistently using one method of instruction and working with only one group of students. To have controlled for the teacher variable, it would have been necessary for each teacher to rotate assignments with high, middle, and low Ss. Had all three groups in each experimental school been exposed to the three different styles of teaching, the results of the present study would not have been subject to teacher bias. (Unfortunately, the fact that each group of students was paced independently of other groups made the rotation of assignments impossible; to have switched groups at any given time would have resulted in some teachers changing assignments in the middle of a lesson.)

Further research is needed to determine the effects of adapting instructional methods to learner characteristics other than entry level of achievement and in curricular areas other than science. Additional studies should also investigate the effects of instructional programming procedures on variables other than achievement and mastery: specifically, study habits, self-concept, aspiration level, motivation, realization of achievement potential, and related personality and social outcomes. More valid assessment of the affective consequences of adapting instruction is needed, as well.

CONCLUSIONS AND IMPLICATIONS

The following major conclusions can be drawn from the results of the present study:

1. In general, grouping for differentiated instruction[2] on the basis of entering achievement levels facilitates performance on tests of science content knowledge and promotes mastery of lesson objectives for students in the high and middle groups, compared with students of similar ability receiving non-differentiated instruction in random-grouped classes.
2. Providing differentiated instruction[2] for students in the low group does not significantly enhance performance on tests of science content knowledge or mastery of lesson objectives, compared with low students receiving non-differentiated instruction in random-grouped classes.
3. No significant differences in attitudes toward science are evident between students receiving differentiated instruction in relatively homogeneous classes and students receiving non-differentiated instruction in relatively heterogeneous classes.

In the present study the Instructional Programming Model could not be fully applied in that the low students were obviously unable to master the content of the difficult SAPA lessons. One recommendation for instruction would be to limit the number of science lessons to be taught to the low ability children and to modify the content by making the concepts less abstract and by using lower level

---

[2]Time and instructional techniques are varied depending upon the needs of the group.

supplementary texts to teach the SAPA content. The implication of the finding that high ability children mastered the SAPA content quickly is that high students could be expected to receive instruction on a greater number of SAPA lessons than provided to middle or low students. In the school system where the present study was conducted, 17 SAPA lessons were identified by the science coordinator and school staffs to be taught at each level (e.g., level six), during a school year. The number of SAPA lessons that have been developed and published for each level actually numbers between 25 and 30.

In the present study the same number of lessons were taught to each group of children; the same objectives were set for every child for the period of time during which the lessons were taught. It was found that the low students were unable to master the content even when instructional time was expanded. This would suggest that merely adapting the time for instruction is not sufficient; the objectives ought to be adapted to the students' entering levels of achievement, as well. Based on the results of the present study, an effective plan for instruction would be the grouping of children according to entering levels of achievement and the adapting of both time and content to be taught. Thus, high students might be expected to master 20 or 25 SAPA lessons per year; middle students should be able to master the usual 17 SAPA lessons; and low students might reasonably be expected to master fewer lessons (i.e., 10 to 15) if the more

difficult concepts are eliminated and the SAPA lessons are supplemented with lower level texts that teach essentially the same content. If time and content are adapted to the entry characteristics of elementary children, it is likely that the resulting instruction will be highly effective.

REFERENCES

Airasian, P. W. An application of a modified version of John Carroll's model of school learning. Unpublished master's thesis, University of Chicago, 1967.

American Association for the Advancement of Science. Science: A process approach. Xerox Corporation, 1967.

Anderson, L. W. Time and school learning. Unpublished doctoral dissertation, University of Chicago, 1973.

Anderson, L. W. An empirical investigation of individual differences in time to learn. Journal of Educational Psychology, 1976, 68, 226-233.

Askov, E. N. Assessment of a system for individualizing reading instruction. Technical Report No. 117. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1970.

Bechtol, W. M., & Conte, A. E. Individually guided social studies. Reading, Mass.: Addison-Wesley, 1976.

Bernard, M. E. The effects of advance organizers and within-text questions on the learning of a taxonomy of concepts. Technical Report No. 357. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Block, J. H. The effects of various levels of performance on selected cognitive, affective, and time variables. Unpublished doctoral dissertation, University of Chicago, 1970.

Block, J. H. (Ed.). Mastery learning: Theory and practice. New York: Holt, Rinehart, and Winston, 1971.

Block, J. H. Student learning and the setting of mastery performance standards. Educational Horizons, 1972, 50, 183-191.

Block, J. H. (Ed.). Schools, society, and mastery learning. New York: Holt, Rinehart, and Winston, 1974.

Block, J. H., & Anderson, L. W. Mastery learning in classroom instruction. New York: Macmillan, 1975.

Block, J. H., & Burns, R. B. Mastery learning. In L. S. Shulman (Ed.), Review of research in education 4. Itasca, Illinois: F. E. Peacock, 1977.

Bloom, B. S. Learning for mastery. UCLA Evaluation Comment, 1968, 1(2), 1-11.

Bloom, B. S. Mastery learning and its implications for curriculum development. In E. W. Eisner (Ed.), Confronting curriculum reform. Boston: Little, Brown and Company, 1971.

Bloom, B. S. An introduction to mastery learning theory. In J. H. Block (Ed.), Schools, society, and mastery learning. New York: Holt, Rinehart, and Winston, 1974.

Bloom, B. S. Human characteristics and school learning. New York: McGraw-Hill, 1976.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

Borg, W. R. Ability grouping in the public schools. Journal of Experimental Education, 1965, 34(2), 1-97.

Box, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems, I: Effect of inequality of variance in the one-way classification. Annals of Mathematical Statistics, 1954, 25, 290-302.

Bruner, J. S. Toward a theory of instruction. Cambridge, Mass.: Harvard University Press, 1966.

Burrows, C. K., & Okey, J. R. The effects of a mastery learning strategy on achievement. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., 1975.

Buss, A. R., & Poley, W. Individual differences: Traits and factors. New York: Gardner Press, 1976.

Carroll, J. A. A model of school learning. Teachers College Record, 1963, 64, 723-733.

Chester, R., Hubbard, W. D., & Van Blaricom, G. Field test report of the self-directed, interpretive, and creative reading elements of the Wisconsin Design for Reading Skill Development. Technical Report No. 330. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Cronbach, L. J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Cronbach, L. J. Comments on mastery learning and its implications for curriculum development. In E. W. Eisner (Ed.), Confronting curriculum reform. Boston: Little, Brown, and Company, 1971.

Dewar, J. A. Grouping for arithmetic instruction in sixth grade. Elementary School Journal, 1963, 63, 266-269.

Dunkin, M. J., & Biddle, B. J. The study of teaching. New York: Holt, Rinehart, and Winston, 1974.

Ekstrom, R. B. Experimental studies of homogeneous grouping: A critical review. School Review, 1961, 69, 216-226.

Esposito, D. Homogeneous and heterogeneous ability groups: Principal findings and implications for evaluating and designing more effective educational environments. Review of Educational Research, 1973, 43, 163-179.

Feldman, K. V. The effects of number of positive and negative instances, concept definition, and emphasis of relevant attributes on the attainment of mathematical concepts. Technical Report No. 243. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1972.

Fiel, R. L., & Okey, J. R. The effects of formative evaluation and remediation on mastery of intellectual skills. The Journal of Educational Research, 1974, 68, 253-255.

Frayer, D. A. Effects of number of instances and emphasis of relevant attribute values on mastery of geometric concepts by fourth- and sixth-grade children. Technical Report No. 116. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1970.

Glaser, R., Reynolds, J. H., & Fullick, M. G. Programmed instruction in the intact classroom. Technical Report. Pittsburgh: University of Pittsburgh Learning Research and Development Center, 1963.

Glaser, R. Adapting the elementary school curriculum to individual performance. In Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N. J.: Educational Testing Service, 1968, Pp. 3-36.

Glass, G. V, Peckham, P. E., & Sanders, J. R. Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. Review of Educational Research, 1972, 42(3), 237-288.

Goldberg, M. L., Passow, A. H., & Justman, J. The effects of ability grouping. New York: Teachers College Press, 1966.

Goodlad, J. I., & Anderson, R. H. The nongraded elementary school. New York: Harcourt, Brace, 1959.

Green, B. A. A self-paced course in freshman physics. Occasional Paper No. 2. Cambridge, Mass.: Educational Research Center, Massachusetts Institute of Technology, 1969.

Haney, R. E., & Sorenson, J. S. Individually guided science. Reading, Mass.: Addison-Wesley, 1977.

Hays, W. L. Statistics for the social sciences (2nd ed.). New York: Holt, Rinehart, and Winston, 1973.

Heathers, G. Grouping. In Encyclopedia of Educational Research (4th ed.). New York: Macmillan, 1969, Pp. 559-570.

Johnson, S. First-year field test of the Wisconsin Design for Reading Skill Development. Technical Report No. 5. Atlanta: Southeastern Educational Laboratory, 1971.

Jones, F. G. The effects of mastery and aptitude on learning, retention, and time. Unpublished doctoral dissertation, University of Georgia, 1974.

Justman, J. Reading and class homogeneity. Reading Teacher, 1968, 21, 314-15, 334.

Katz, S. The effect of each of four instructional treatments on the learning of principles by children. Technical Report No. 381. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Keller, F. S. Goodbye, teacher . . . Journal of Applied Behavior Analysis, 1968, 1, 79-89.

Kersh, M. E. A strategy for mastery learning in fifth-grade arithmetic. Unpublished doctoral dissertation, University of Chicago, 1971.

Kim, H. Mastery learning in the middle school. Seoul, Korea: Korean Institute for Research in the Behavioral Sciences, 1970.

Kim, H. Mastery learning in the Korean middle schools. UNESCO Regional Office for Education in Asia, September 1971, 6(1), Sec. I, 55-66.

Kim, Y., Cho, G., Park, J., & Park, N. An application of a new instructional model. Research Report No. 8. Seoul, Korea: Korean Educational Development Institute, 1974.

King, I. L., & Romberg, T. A. A formative development of a unit on proof for use in the elementary school. Technical Report No. 111. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1970.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole Publishing Co., 1968.

Klausmeier, H. J., & Goodwin, W. Learning and human abilities (4th ed.). New York: Harper & Row, 1975.

Klausmeier, H. J. Conceptual development during the school years. In J. R. Levin & V. L. Allen (Eds.), Cognitive learning in children: Theories and strategies. New York: Academic Press, 1976.

Klausmeier, H. J. Instructional programming for the individual student. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Klausmeier, H. J., Allen, P. S., Sipple, T. S., & White, K. M. First longitudinal study of attainment of the concept equilateral triangle by children age 5 to 16. Working Paper No. 163. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Klausmeier, H. J., DiLuzio, G. L., & Brumet, M. L. First year of intervention study to facilitate children's concept learning. Working Paper No. 159. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Klausmeier, H. J., & Feldman, K. V. The effects of a definition and and a varying number of examples and nonexamples on concept attainment. Journal of Educational Psychology, 1975, 67, 174-178.

Klausmeier, H. J., Jeter, J. T., Quilling, M. R., Frayer, D. A., & Allen, P. A. Individually guided motivation. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Klausmeier, H. J., Quilling, M. R., Sorenson, J. S., Way, R. S., & Glasrud, G. R. Individually guided education and the multiunit school: Guidelines for implementation. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1971.

Klausmeier, H. J., Schilling, J. M., & Feldman, K. V. The effectiveness of instructional lessons in accelerating children's attainment of the concept tree. Technical Report No. 372. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Klausmeier, H. J., Sipple, T. S., & Allen, P. S. First cross-sectional study of attainment of the concepts equilateral triangle, cutting tool, and noun by children age 5 to 16 of city A. Technical Report No. 287. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1974.

Klausmeier, H. J., Sipple, T. S., Swanson, J. E., & Schilling, J. M. Second-year intervention study to facilitate children's concept learning. Working Paper No. 193. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Klausmeier, H. J., Sorenson, J. S., & Quilling, M. R. Instructional programming for the individual pupil in the multiunit elementary school. Elementary School Journal, 1971, 72, 88-101.

Klumb, R. W. The effects of monitoring pupil performance and two incentive treatments for teachers on pupils' reading skill development and teachers' attitudes. Technical Report No. 268. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1973.

Lee, Y. D., and others. Interaction improvement studies on the mastery learning project--Final report on mastery learning program, April-November 1971. Seoul, Korea: Educational Research Center, Seoul National University, 1971.

McMurray, N. E., Bernard, M. E., & Klausmeier, H. J. An instructional design for accelerating children's concept learning. Technical Report No. 321. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Morrison, H. C. The practice of teaching in the secondary school. Chicago: University of Chicago Press, 1926.

Moser, J. M., & Romberg, T. A. The evaluation of DMP. Project Paper 77-2. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1977.

Mueller, D. J. Mastery learning: Partly boon, partly boondoggle. Teachers College Record, 1976, 78(1), 41-52.

Okey, J. R. Altering teacher and pupil behavior with mastery teaching. School Science and Mathematics, 1974, 74, 530-535.

Okey, J. R. Development of mastery teaching materials. Final Evaluation Report, USOE G-74-2990, Indiana University, 1975.

Otto, W. The Wisconsin Design: A reading program for individually guided education. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Otto, W., & Chester, R. Objective-based reading. Reading, Mass.: Addison-Wesley, 1976.

Peterson, P. L. A review of the research on mastery-learning strategies. Unpublished manuscript, International Association for the Evaluation of Educational Achievement, Stockholm, Sweden, 1972.

Poggio, J. P., Glasnapp, D. R., & Ory, J. C. The impact of test anxiety on formative and summative exam performance in the mastery learning model. Paper presented at the annual meeting of

the National Council on Measurement in Education, Washington, D.C., 1975.

Quilling, M. R. The first field test of the Word Attack element of the Wisconsin Design for Reading Skill Development. Technical Memo No. QV-10-70. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1971.

Quilling, M. R., & Otto, W. Evaluation of an objective based curriculum in reading. Journal of Educational Research, 1971, 65, 15-18.

Quilling, M. R., & Wojtal, P. The Wisconsin Design for Reading Skill Development: Study Skills--A report on the pilot test: 1970-72. Working Paper No. 97. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1972.

Romberg, T. A. Individually guided mathematics. Reading, Mass.: Addison-Wesley, 1976.

Romberg, T. A. Developing mathematical processes: The elementary mathematics program for individually guided education. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Romberg, T. A., Shepler, J. L., & King, I. L. Mastery learning and retention. Technical Report No. 151. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1970.

Rude, R. T. Implementing and field-testing of the Design. In F. P. Green (Ed.), Investigations relating to mature reading, Twenty-first Yearbook of the National Reading Conference. Milwaukee: National Reading Conference, Inc., 1972, Pp. 265-272.

Rude, R. T. Sex, intelligence and school reading curriculum as factors influencing summer retention of overall reading ability and specific reading skills of first grade subjects. Technical Report No. 263. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1973.

Sals, D., & Hubbard, W. D. The Wisconsin Design for Reading Skill Development: Study Skills--A report on the Type I field test, 1971-73. Technical Report No. 341. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1975.

Scheffé, H. The analysis of variance. New York: Wiley, 1959.

Schilling, J. M. The effects of specially designed concept lessons and instructional programming for the individual student on children's achievement in science. Working Paper No. 182. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Schilling, J. M., Bull, K. S., Feldman, K. V., & Klausmeier, H. J. A method for determining the difficulty of concept instances, taxonomic items, and vocabulary items. Working Paper No. 174. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Shepler, J. L., & Romberg, T. A. A study of parts of the development of a unit in probability and statistics for the elementary school. Technical Report No. 105. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1969.

Skinner, B. F. The science of learning and the art of teaching. Harvard Educational Review, 1954, 24(2), 86-97.

Suppes, P. Modern learning theory and the elementary school curriculum. American Educational Research Journal, 1964, 1, 79-93.

Suppes, P. The uses of computers in education. Scientific American, 1966, 215(51), 206-221.

Swanson, J. E. The effects of number of positive and negative instances, concept definition, and emphasis of relevant attributes on the attainment of three environmental concepts by sixth-grade children. Technical Report No. 244. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1972.

Tennyson, R. D., Woolley, F. R., & Merrill, M. D. Exemplar and non-exemplar variables which produce correct classification behavior and specified classification errors. Journal of Educational Psychology, 1972, 63, 144-152.

Thompson, R. B. Diagnosis and remedial instruction in mathematics. School Science and Mathematics, 1941, 41, 125-128.

Venezky, R. L., & Pittelman, S. D. PRS: A pre-reading skills program for individually guided education. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided

elementary education: Concepts and practices. New York: Academic Press, 1977.

Washburne, C. Adjusting the school to the child. New York: World Book Company, 1932.

Wearne, D. C., & Whitaker, D. R. The 1976 DMP evaluation report. Working Paper No. 185. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

Willcutt, R. E. Ability grouping by content topics in junior high mathematics. Journal of Educational Research, 1969, 63, 152-156.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Wisconsin Research and Development Center for Cognitive Learning. Individualized schooling. Technical Proposal: 1977-1979. Madison, Wisconsin, 1976.

Yeager, J. L., & Kissel, M. A. An investigation of the relationship between selected student characteristics and time required to achieve mastery. Working Paper No. 46. Pittsburgh: Learning Research and Development Center, University of Pittsburgh, 1969.

Appendix A has been omitted from this publication, but is available on microfilm from Xerox University Microfilms, Ann Arbor, Michigan 48106

APPENDIX B

NOTES TO TEACHERS AND SUMMARY OF INSTRUCTIONAL PROCEDURES TO BE CARRIED OUT IN EXPERIMENTAL AND CONTROL SCHOOLS

NOTES TO EXPERIMENTAL TEACHERS AND SUMMARY OF INSTRUCTIONAL PROCEDURES TO BE CARRIED OUT IN EXPERIMENTAL SCHOOLS

Teachers: Here are some important points to keep in mind regarding instruction and testing in science for the remainder of this semester.

A. The major objective is to move through the SAPA lessons in the prescribed sequence at a pace that gives most students the opportunity to master the content.

B. Each teacher will teach six SAPA lessons during the next eight to ten weeks to one group of children (high, middle, or low). How quickly the students move from one lesson to the next will depend upon the needs of the group.

C. The following sequence of instruction and testing should be strictly observed in the experimental schools:

1. Provide initial instruction on a SAPA lesson using the methods and materials described in the teacher guide; try to employ a variety of large-group, small-group, and individual activities.

2. When most (approximately four-fifths) of the students understand the lesson content, administer the 15-item diagnostic test. Read all items aloud and give ample time for children to mark their responses on the answer sheets.

3. Give the answer sheets to the instructional aide for immediate grading.

4. Return the graded answer sheets to the students and review each test item that was missed by three or more students.

5. Use the diagnostic tests to determine where problems occurred and give extra instruction on these weak spots. You may wish to separate those students achieving mastery (80% correct) from those not attaining mastery, and provide additional instruction for the non-mastery group. Another possibility is to have the mastery students tutor the non-mastery students. Try having some students who understand the content to act as "teachers" for certain items that were missed. Whether you use limited tutoring or choose to do the teaching yourself, provide for coverage of the most commonly missed items and try to re-teach those children who did not master the material.

6. After re-teaching, do not provide general enrichment or supplementary activities on the SAPA lesson, but move immediately to the next lesson in the sequence. Following the completion of all six SAPA lessons, enrichment may be provided if time allows. Once again, the object is to bring children to mastery on one lesson and move on to the next.

D. Please keep a daily log of class activities. Indicate the number of minutes spent, the type of activity (initial teaching, testing, re-teaching) and the structure of the class (large group, small group, individual). Also record any personal observations or comments which would characterize your class, and mention specific problems (if any) that are encountered.

NOTES TO CONTROL TEACHERS AND SUMMARY OF INSTRUCTIONAL PROCEDURES TO BE CARRIED OUT IN CONTROL SCHOOLS

Teachers: Here are some important points to keep in mind regarding instruction and testing in science for the remainder of the semester.

A. The major objective is to present the SAPA material in a way that will allow as many children as possible to master the content.

B. Each teacher will teach six SAPA lessons during the next eight to ten weeks to classes of heterogeneously grouped children.

C. The following sequence of instruction and testing should be strictly observed in the control schools:

1. Provide initial instruction on a SAPA lesson using the methods and materials described in the teacher guide.

2. When most of the students understand the lesson content, administer the 15-item SAPA test. Read all items aloud and give ample time for children to mark their responses on the answer sheets.

3. Collect all tests and answer sheets when students have finished; do not discuss specific test questions, and do not provide a formal review session pertaining to the test items.

4. Following the SAPA test move immediately to the next SAPA lesson to be taught. Do not provide enrichment or supplementary activities until all six SAPA lessons have been taught.

5. Please keep an accurate account of the number of minutes spent on instruction for each SAPA lesson. Also indicate what activities were covered and describe the structure of the class (large group, small group, individual). Please record any personal observations or comments which would characterize your class, and mention specific problems (if any) that are encountered.

APPENDIX C

SCIENCE ATTITUDE INVENTORY

Name _____ Level _____
(last) (first)

School _____ Teacher _____

	Agree	Slightly Agree	Slightly Disagree	Disagree
1. Science is a waste of time.				
2. I can hardly wait for the school day to be over.				
3. Science is interesting.				
4. I would like to get a microscope or chemistry set for Christmas or my birthday.				
5. Science is a good hobby.				
6. I think science classes are too long.				
7. I like anything that has to do with science.				
8. I wish we spent more time studying plants and animals.				
9. I like science books more than any other kind of books.				
10. I wish that history books were more interesting.				

	Agree	Slightly Agree	Slightly Disagree	Disagree
11. I could do very well without science.				
12. I wish that I could write better stories.				
13. Science is a good subject.				
14. I would like to learn to speak a foreign language well.				
15. I am not interested in science.				
16. I would enjoy going to school during the summer if I could study science.				
17. If it has to do with science, forget it! I hate science.				
18. I would like to be a scientist when I grow up.				
19. I like to read the newspaper.				
20. There are lots of things I'd rather do than study science.				
21. I enjoy math better than spelling.				
22. I have no desire to study science.				
23. I can hardly wait for math class.				
24. I only take science because I have to.				

APPENDIX D

PRETEST NUMBERS, MEANS AND STANDARD DEVIATIONS FOR EACH SAPA-CONTENT TEST

PRETEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 10-ITEM SAPA-CONTENT TEST L (GUINEA PIGS IN A MAZE)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	8.00	7.92	8.25	7.78	8.14	7.84
S.D.	1.84	1.38	1.07	1.22	1.46	1.27
Middle						
Number	20	16	18	24	38	40
Mean	6.65	6.38	6.17	5.58	6.42	5.90
S.D.	1.18	1.71	1.62	1.74	1.41	1.75
Low						
Number	20	27	23	12	43	39
Mean	4.20	4.52	4.17	4.92	4.19	4.64
S.D.	1.44	2.08	2.21	2.07	1.87	2.06
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	6.28	5.84	6.23	6.17	6.26	6.00
S.D.	2.17	2.29	2.42	2.02	2.29	2.15

PRETEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 10-ITEM SAPA-CONTENT TEST M (OBSERVATIONS AND HYPOTHESES)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	7.80	5.77	8.33	7.22	8.09	6.61
S.D.	1.11	2.17	1.58	1.52	1.39	1.93
Middle						
Number	20	16	18	24	38	40
Mean	5.15	4.25	5.61	5.25	5.37	4.85
S.D.	1.46	1.06	2.00	1.33	1.73	1.31
Low						
Number	20	27	23	12	43	39
Mean	4.70	3.93	3.57	4.00	4.09	3.95
S.D.	1.38	1.03	2.00	1.48	1.81	1.17
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	5.88	4.47	5.89	5.63	5.89	5.03
S.D.	1.90	1.55	2.75	1.87	2.37	1.80

PRETEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 10-ITEM SAPA-CONTENT TEST N (SUFFOCATING CANDLE)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	5.80	5.62	6.25	5.89	6.05	5.77
S.D.	1.51	1.04	1.54	2.03	1.52	1.67
Middle						
Number	20	16	18	24	38	40
Mean	4.75	4.19	4.22	5.21	4.50	4.80
S.D.	1.21	1.11	1.90	1.64	1.57	1.52
Low						
Number	20	27	23	12	43	39
Mean	4.05	3.22	3.00	4.00	3.49	3.46
S.D.	2.21	1.65	1.45	1.21	1.89	1.55
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	4.87	4.05	4.54	5.17	4.70	4.60
S.D.	1.82	1.67	2.12	1.81	1.98	1.82

PRETEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 10-ITEM SAPA-CONTENT TEST O (USING PUNCH CARDS)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	4.20	4.62	4.54	4.83	4.39	4.74
S.D.	1.51	1.76	1.50	1.62	1.50	1.65
Middle						
Number	20	16	18	24	38	40
Mean	4.15	3.44	4.33	4.08	4.24	3.83
S.D.	1.53	1.26	1.33	1.47	1.42	1.41
Low						
Number	20	27	23	12	43	39
Mean	2.65	3.37	2.96	3.17	2.81	3.31
S.D.	1.39	1.31	1.64	1.80	1.52	1.45
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	3.67	3.68	3.92	4.13	3.80	3.90
S.D.	1.62	1.48	1.65	1.68	1.64	1.59

PRETEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 10-ITEM SAPA-CONTENT TEST P (DESCRIBING LOCATION)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	6.45	5.92	6.71	5.11	6.59	5.45
S.D.	1.43	1.80	1.46	1.75	1.44	1.79
Middle						
Number	20	16	18	24	38	40
Mean	4.70	5.63	4.56	4.79	4.63	5.13
S.D.	1.95	1.15	1.25	1.89	1.63	1.67
Low						
Number	20	27	23	12	43	39
Mean	3.50	3.04	2.91	2.92	3.19	3.00
S.D.	1.67	1.60	1.59	1.62	1.64	1.59
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	4.88	4.45	4.77	4.48	4.82	4.46
S.D.	2.07	2.04	2.17	1.95	2.11	1.99

PRETEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 10-ITEM SAPA-CONTENT TEST Q (REPORTING AN INVESTIGATION)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	7.45	7.23	7.25	6.33	7.34	6.71
S.D.	1.32	1.83	1.94	1.19	1.67	1.53
Middle						
Number	20	16	18	24	38	40
Mean	6.05	5.88	6.11	5.71	6.08	5.78
S.D.	1.42	1.45	1.71	1.08	1.75	1.23
Low						
Number	20	27	23	12	43	39
Mean	3.95	3.63	3.83	3.58	3.88	3.62
S.D.	1.61	1.28	1.15	1.78	1.37	1.43
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	5.82	5.11	5.72	5.44	5.77	5.27
S.D.	2.14	2.09	2.19	1.64	2.16	1.89

APPENDIX E

POSTTEST NUMBERS, MEANS AND STANDARD DEVIATIONS FOR EACH SAPA-CONTENT TEST

POSTTEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 15-ITEM SAPA-CONTENT TEST L (GUINEA PIGS IN A MAZE)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	13.55	13.46	13.83	11.33	13.70	12.23
S.D.	1.15	1.20	1.09	1.57	1.11	1.76
Middle						
Number	20	16	18	24	38	40
Mean	12.40	12.31	13.56	10.21	12.95	11.05
S.D.	1.47	1.35	0.70	2.62	1.29	2.42
Low						
Number	20	27	23	12	43	39
Mean	9.90	10.63	9.83	8.58	9.86	10.00
S.D.	3.23	2.15	3.92	2.50	3.57	2.43
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	11.95	11.77	12.34	10.22	12.15	11.01
S.D.	2.61	2.10	3.06	2.47	2.85	2.41

POSTTEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 15-ITEM SAPA-CONTENT TEST M (OBSERVATIONS AND HYPOTHESES)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	14.05	11.62	14.04	10.83	14.05	11.16
S.D.	1.00	2.57	1.20	2.87	1.10	2.73
Middle						
Number	20	16	18	24	38	40
Mean	12.30	10.38	12.33	9.25	12.32	9.70
S.D.	1.75	2.31	2.54	2.38	2.13	2.39
Low						
Number	20	27	23	12	43	39
Mean	7.50	8.67	8.13	8.67	7.84	8.67
S.D.	2.84	2.20	3.45	2.23	3.16	2.18
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	11.28	9.84	11.48	9.65	11.38	9.75
S.D.	3.42	2.59	3.61	2.63	3.51	2.60

POSTTEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 15-ITEM SAPA-CONTENT TEST N (SUFFOCATING CANDLE)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	13.00	11.85	14.13	10.06	13.61	10.81
S.D.	2.22	2.54	1.26	2.29	1.83	2.52
Middle						
Number	20	16	18	24	38	40
Mean	12.90	10.50	11.00	8.42	12.00	9.25
S.D.	1.86	1.71	2.09	1.82	2.17	2.03
Low						
Number	20	27	23	12	43	39
Mean	7.55	8.85	8.57	7.17	8.09	8.33
S.D.	3.00	2.38	2.98	1.85	3.00	2.34
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	11.15	10.02	11.29	8.69	11.22	9.36
S.D.	3.49	2.53	3.24	2.24	3.35	2.47

POSTTEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 15-ITEM SAPA-CONTENT TEST O (USING PUNCH CARDS)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	13.60	13.00	13.75	11.41	13.68	12.08
S.D.	1.35	1.73	1.59	1.78	1.47	1.91
Middle						
Number	20	16	18	24	38	40
Mean	13.25	11.69	12.56	10.33	12.92	10.88
S.D.	1.80	2.63	1.62	2.20	1.73	2.44
Low						
Number	20	27	23	12	43	39
Mean	8.85	11.78	10.09	8.67	9.51	10.82
S.D.	1.98	2.04	3.41	2.57	2.87	2.62
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	11.90	12.04	12.12	10.32	12.02	11.19
S.D.	2.77	2.19	2.85	2.35	2.80	2.42

POSTTEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 15-ITEM SAPA-CONTENT TEST P (DESCRIBING LOCATION)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	13.30	14.00	13.46	11.00	13.39	12.26
S.D.	1.69	1.41	1.59	2.83	1.62	2.76
Middle						
Number	20	16	18	24	38	40
Mean	13.30	13.06	11.33	10.46	12.37	11.50
S.D.	1.69	1.39	2.74	2.38	2.43	2.40
Low						
Number	20	27	23	12	43	39
Mean	10.90	10.19	10.78	9.42	10.84	9.95
S.D.	2.00	2.50	3.12	3.00	2.63	2.65
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	12.50	11.89	11.92	10.41	12.20	11.16
S.D.	2.10	2.61	2.77	2.69	2.48	2.74

POSTTEST NUMBERS, UNADJUSTED MEANS, AND STANDARD DEVIATIONS BY GROUP, FOR EXPERIMENTAL AND CONTROL SCHOOLS: 15-ITEM SAPA-CONTENT TEST Q (REPORTING AN INVESTIGATION)

	School Comparison					
Group	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High						
Number	20	13	24	18	44	31
Mean	12.70	12.00	12.71	10.83	12.70	11.32
S.D.	1.03	1.96	1.04	2.60	1.02	2.39
Middle						
Number	20	16	18	24	38	40
Mean	10.75	11.50	11.61	9.46	11.16	10.28
S.D.	1.94	2.22	1.54	2.77	1.79	2.73
Low						
Number	20	27	23	12	43	39
Mean	7.35	8.07	8.22	7.25	7.81	7.82
S.D.	1.79	3.16	3.15	3.05	2.61	3.11
Combined Across Groups						
Number	60	56	65	54	125	110
Mean	10.27	9.96	10.82	9.43	10.55	9.70
S.D.	2.75	3.21	2.89	3.03	2.83	3.12

APPENDIX F

NUMBERS OF STUDENTS ATTAINING MASTERY

FOR EACH SAPA-CONTENT LESSON

NUMBERS OF STUDENTS AT MASTERY AND NON-MASTERY AND

RESULTS OF CHI-SQUARE TESTS BY GROUP:

SAPA-CONTENT POSTTEST L (GUINEA PIGS IN A MAZE)

High Group

	E_1	C_1
Mastery	19	12
Non-Mastery	1	1

$\chi^2 = 0.19$

	E_2	C_2
Mastery	24	11
Non-Mastery	0	7

$\chi^2 = 8.58^{**}$

	E_1E_2	C_1C_2
Mastery	43	23
Non-Mastery	1	8

$\chi^2 = 7.44^{**}$

Middle Group

	E_1	C_1
Mastery	15	10
Non-Mastery	5	6

$\chi^2 = 0.20$

	E_2	C_2
Mastery	18	11
Non-Mastery	0	13

$\chi^2 = 11.70^{***}$

	E_1E_2	C_1C_2
Mastery	33	21
Non-Mastery	5	19

$\chi^2 = 9.24^{**}$

Low Group

	E_1	C_1
Mastery	8	10
Non-Mastery	12	6

$\chi^2 = 1.01$

	E_2	C_2
Mastery	8	1
Non-Mastery	15	11

$\chi^2 = 1.67$

	E_1E_2	C_1C_2
Mastery	16	11
Non-Mastery	27	17

$\chi^2 = 0.01$

Combined Groups Within School

	E_1	C_1
Mastery	42	31
Non-Mastery	18	25

$\chi^2 = 2.07$

	E_2	C_2
Mastery	50	23
Non-Mastery	15	31

$\chi^2 = 13.25^{***}$

	E_1E_2	C_1C_2
Mastery	92	54
Non-Mastery	33	56

$\chi^2 = 13.91^{***}$

$^{*}p < .05$
$^{**}p < .01$
$^{***}p < .001$

NUMBERS OF STUDENTS AT MASTERY AND NON-MASTERY AND

RESULTS OF CHI-SQUARE TESTS BY GROUP:

SAPA-CONTENT POSTTEST M (OBSERVATIONS & HYPOTHESES)

High Group

	E_1	C_1
Mastery	20	8
Non-Mastery	0	5

$\chi^2 = 6.32^*$

	E_2	C_2
Mastery	22	11
Non-Mastery	2	7

$\chi^2 = 4.03^*$

	E_1E_2	C_1C_2
Mastery	42	19
Non-Mastery	2	12

$\chi^2 = 11.82^{***}$

Middle Group

	E_1	C_1
Mastery	13	6
Non-Mastery	7	10

$\chi^2 = 1.71$

	E_2	C_2
Mastery	12	3
Non-Mastery	6	21

$\chi^2 = 10.89^{***}$

	E_1E_2	C_1C_2
Mastery	25	9
Non-Mastery	13	31

$\chi^2 = 13.14^{***}$

Low Group

	E_1	C_1
Mastery	1	5
Non-Mastery	19	22

$\chi^2 = 0.87$

	E_2	C_2
Mastery	5	0
Non-Mastery	18	12

$\chi^2 = 1.53$

	E_1E_2	C_1C_2
Mastery	6	5
Non-Mastery	37	34

$\chi^2 = 0.03$

Combined Groups Within School

	E_1	C_1
Mastery	34	19
Non-Mastery	26	37

$\chi^2 = 5.15^*$

	E_2	C_2
Mastery	39	14
Non-Mastery	26	40

$\chi^2 = 12.52^{***}$

	E_1E_2	C_1C_2
Mastery	73	33
Non-Mastery	52	77

$\chi^2 = 17.93^{***}$

$^*p < .05$
$^{**}p < .01$
$^{***}p < .001$

NUMBERS OF STUDENTS AT MASTERY AND NON-MASTERY AND

RESULTS OF CHI-SQUARE TESTS BY GROUP:

SAPA-CONTENT POSTTEST N (SUFFOCATING CANDLE)

High Group

	E_1	C_1
Mastery	14	7
Non-Mastery	6	6

$\chi^2 = 0.33$

	E_2	C_2
Mastery	23	5
Non-Mastery	1	13

$\chi^2 = 18.48^{***}$

	E_1E_2	C_1C_2
Mastery	37	12
Non-Mastery	7	19

$\chi^2 = 14.59^{***}$

Middle Group

	E_1	C_1
Mastery	17	5
Non-Mastery	3	11

$\chi^2 = 8.66^{**}$

	E_2	C_2
Mastery	8	1
Non-Mastery	10	23

$\chi^2 = 7.66^{**}$

	E_1E_2	C_1C_2
Mastery	25	6
Non-Mastery	13	34

$\chi^2 = 18.92^{***}$

Low Group

	E_1	C_1
Mastery	2	3
Non-Mastery	18	24

$\chi^2 = 0.13$

	E_2	C_2
Mastery	5	0
Non-Mastery	18	12

$\chi^2 = 1.53$

	E_1E_2	C_1C_2
Mastery	7	3
Non-Mastery	36	36

$\chi^2 = 0.72$

Combined Groups Within School

	E_1	C_1
Mastery	33	15
Non-Mastery	27	41

$\chi^2 = 8.38^{**}$

	E_2	C_2
Mastery	36	6
Non-Mastery	29	48

$\chi^2 = 23.41^{***}$

	E_1E_2	C_1C_2
Mastery	69	21
Non-Mastery	56	89

$\chi^2 = 30.77^{***}$

$^{*}p < .05$
$^{**}p < .01$
$^{***}p < .001$

NUMBERS OF STUDENTS AT MASTERY AND NON-MASTERY AND

RESULTS OF CHI-SQUARE TESTS BY GROUP:

SAPA-CONTENT POSTTEST 0 (USING PUNCH CARDS)

	E_1	C_1	E_2	C_2	E_1E_2	C_1C_2
High Group						
Mastery	18	11	22	11	40	22
Non-Mastery	2	2	2	6	4	8
	$\chi^2 = 0.01$		$\chi^2 = 3.05$		$\chi^2 = 2.87$	
Middle Group						
Mastery	17	9	14	8	31	17
Non-Mastery	3	7	4	16	7	23
	$\chi^2 = 2.37$		$\chi^2 = 6.46^*$		$\chi^2 = 10.98^{***}$	
Low Group						
Mastery	2	14	10	1	12	15
Non-Mastery	18	13	13	11	31	24
	$\chi^2 = 7.20^{**}$		$\chi^2 = 3.04$		$\chi^2 = 0.61$	
Combined Groups Within School						
Mastery	37	34	46	20	83	54
Non-Mastery	23	22	19	33	42	55
	$\chi^2 = 0.01$		$\chi^2 = 11.62^{***}$		$\chi^2 = 6.14^*$	

$^*p < .05$
$^{**}p < .01$
$^{***}p < .001$

NUMBERS OF STUDENTS AT MASTERY AND NON-MASTERY AND RESULTS OF CHI-SQUARE TESTS BY GROUP: SAPA-CONTENT POSTTEST P (DESCRIBING LOCATION)

		E_1	C_1		E_2	C_2		E_1E_2	C_1C_2
High Group	Mastery	16	12	Mastery	21	8	Mastery	37	20
	Non-Mastery	4	1	Non-Mastery	3	10	Non-Mastery	7	11
		$\chi^2 = 0.22$			$\chi^2 = 7.02^{**}$			$\chi^2 = 2.82$	
Middle Group	Mastery	17	14	Mastery	9	10	Mastery	26	24
	Non-Mastery	3	2	Non-Mastery	9	14	Non-Mastery	12	16
		$\chi^2 = 0.07$			$\chi^2 = 0.05$			$\chi^2 = 0.29$	
Low Group	Mastery	9	7	Mastery	11	4	Mastery	20	11
	Non-Mastery	11	20	Non-Mastery	12	8	Non-Mastery	23	28
		$\chi^2 = 1.11$			$\chi^2 = 0.21$			$\chi^2 = 2.19$	
Combined Groups Within School	Mastery	42	33	Mastery	41	22	Mastery	83	55
	Non-Mastery	18	23	Non-Mastery	24	32	Non-Mastery	42	55
		$\chi^2 = 1.11$			$\chi^2 = 5.04^{*}$			$\chi^2 = 5.83^{*}$	

*$p < .05$
**$p < .01$
***$p < .001$

NUMBERS OF STUDENTS AT MASTERY AND NON-MASTERY AND

RESULTS OF CHI-SQUARE TESTS BY GROUP:

SAPA-CONTENT POSTTEST Q (REPORTING AN INVESTIGATION)

High Group

	E_1	C_1
Mastery	17	9
Non-Mastery	3	4

$\chi^2 = 0.42$

	E_2	C_2
Mastery	22	9
Non-Mastery	2	9

$\chi^2 = 7.21^{**}$

	E_1E_2	C_1C_2
Mastery	39	18
Non-Mastery	5	13

$\chi^2 = 7.72^{**}$

Middle Group

	E_1	C_1
Mastery	10	9
Non-Mastery	10	7

$\chi^2 = 0.003$

	E_2	C_2
Mastery	11	7
Non-Mastery	7	17

$\chi^2 = 3.08$

	E_1E_2	C_1C_2
Mastery	21	16
Non-Mastery	17	24

$\chi^2 = 1.26$

Low Group

	E_1	C_1
Mastery	0	3
Non-Mastery	20	24

$\chi^2 = 0.88$

	E_2	C_2
Mastery	4	1
Non-Mastery	19	11

$\chi^2 = 0.05$

	E_1E_2	C_1C_2
Mastery	4	4
Non-Mastery	39	35

$\chi^2 = 0.05$

Combined Groups Within School

	E_1	C_1
Mastery	27	21
Non-Mastery	33	35

$\chi^2 = 0.40$

	E_2	C_2
Mastery	37	17
Non-Mastery	28	37

$\chi^2 = 6.71^{**}$

	E_1E_2	C_1C_2
Mastery	64	38
Non-Mastery	61	72

$\chi^2 = 5.95^{*}$

$^{*}p < .05$
$^{**}p < .01$
$^{***}p < .001$

Center Planning and Policy Committee

Richard A. Rossmiller
Wayne Otto
Center Co-directors

Wayne Otto
Area Chairperson
Studies in Reading, Language and Communication

Marvin J. Fruth
Area Chairperson
Studies of Implementation of Individualized Schooling

Herbert J. Klausmeier
Area Chairperson
Studies of Instructional Programming for the Individual Student

James M. Lipham
Area Chairperson
Studies of Administration and Organization for Instruction

Thomas A. Romberg
Area Chairperson
Studies in Mathematics and Evaluation of Practices in Individualized Schooling

Associated Faculty

Vernon L. Allen
Professor
Psychology

B. Dean Bowles
Professor
Educational Administration

Thomas P. Carpenter
Associate Professor
Curriculum and Instruction

W. Patrick Dickson
Assistant Professor
Child and Family Studies

Marvin J. Fruth
Professor
Educational Administration

John G. Harvey
Professor
Mathematics
Curriculum and Instruction

Frank H. Hooper
Professor
Child and Family Studies

Dale D. Johnson
Professor
Curriculum and Instruction

Herbert J. Klausmeier
V.A.C. Henmon Professor
Educational Psychology

Joseph T. Lawton
Assistant Professor
Child and Family Studies

Joel R. Levin
Professor
Educational Psychology

James M. Lipham
Professor
Educational Administration

Dominic W. Massaro
Professor
Psychology

Donald M. McIsaac
Professor
Educational Administration

Wayne R. Otto
Professor
Curriculum and Instruction

Penelope L. Peterson
Assistant Professor
Educational Psychology

Robert G. Petzold
Professor
Music
Curriculum and Instruction

Thomas S. Popkewitz
Associate Professor
Curriculum and Instruction

Gary G. Price
Assistant Professor
Curriculum and Instruction

Thomas A. Romberg
Professor
Curriculum and Instruction

Richard A. Rossmiller
Professor
Educational Administration

B. Robert Tabachnick
Professor
Curriculum and Instruction

J. Fred Weaver
Professor
Curriculum and Instruction

Gary G. Wehlage
Associate Professor
Curriculum and Instruction

END 192